

MEBLINE	20530913
PUMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)



[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK
REFERENCE	(bases 1 to 1684)			
AUTHORS	Li, W.-B., Gruber, C., Tessee, J. and Polayes, D.			
TITLE	Full-length cDNA libraries and normalization			
JOURNAL	Unpublished			
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 1684)			
REFERENCE	Genoscope.			
AUTHORS	Genoscope.			
TITLE	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :			
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr			
COMMENT	- Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.			
FEATURES	Location/Qualifiers			
SOURCE	1..1684			
ORIGIN	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DD001YCO5" /tissue_type="Neuroblastoma Cot 50-normalized" /plasmid="pCMVSPORT_6"			
Query Match	74.6% ; Score 1423 ; DB 3 ; Length 1684 ;			
Best Local Similarity	100.0% ; Pred. No.5.4e-273 ;			
Matches 1423 ;	Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;			
QY	486 GGCCGGCAGTCGGCGCCATGCTGCTGCCCTGCACACCCCTCAACCTGGACCATACCTTATTTGA 545			
DB	1 GGCCGGCAGTCGGCGCCATGCTGCTGCCCTGCACACCCCTCAACCTGGACCATACCTTATTTGA 60			
QY	546 CGCAGTCGCCCCAGGCGGCTTTGCCGACGCTGGTCAAGCTCTCCGCTGGAAGCTCAAGCTC 605			
DB	61 CGCAGTCGCCCCAGGCGGCTTTGCCGACGCTGGTCAAGCTCTCCGCTGGAAGCTCAAGCTC 120			
QY	606 CAACCGGCTGGGCGACGCTGGGCTCCGGACCGGCTTTCTCTGTTGGGGGTATGAGAGAGC 665			
DB	121 CAACCGGCTGGGCGACGCTGGGCTCCGGACCGGCTTTCTCTGTTGGGGGTATGAGAGAGC 180			
QY	666 CTCTCCGCGCCCTCTGGTGTGAGCTTTAGCGGGAAACCCCTGCATGCAACTGTGAGCT 725			
DB	181 CTCTCCGCGCCCTCTGGTGTGAGCTTTAGCGGGAAACCCCTGCATGCAACTGTGAGCT 240			
QY	726 GCTGTGGCTGGGCGGCTGGGCGGCGCGGACGACCTTGAAACGTTGGGCTTCCGCGCCGG 785			
DB	241 GCTGTGGCTGGGCGGCTGGGCGGCGCGGACGACCTTGAAACGTTGGGCTTCCGCGCCGG 300			
QY	786 CTTGGCGCGCGGCTACTTCTGGGCGAGGCGCGAGGGGAGATTTCTCTGTGAGCGCGCCCT 845			
DB	301 CTTGGCGCGCGGCTACTTCTGGGCGAGGCGCGAGGGGAGATTTCTCTGTGAGCGCGCCCT 360			
QY	846 CATTGCGCGCGCACAGCAGCGGCTTGGTGGTCTGTAAGAGCGAGCGGCGCACAGCTGCGGTTG 905			
DB	361 CATTGCGCGCGCACAGCAGCGGCTTGGTGGTCTGTAAGAGCGAGCGGCGCACAGCTGCGGTTG 420			
QY	906 CCGGCGCTTGGGTGACCCCGCGGCTTACCATGACACTGGGTGGTCTTACGACGCGTTGGT 965			
DB	421 CCGGCGCTTGGGTGACCCCGCGGCTTACCATGACACTGGGTGGTCTTACGACGCGTTGGT 480			
QY	966 TGGCAACTCTCCCGGACCGCGGCTTCCGCAACGGGACCTTGAAGATTGGGGGTGACCGG 1025			
DB	481 TGGCAACTCTCCCGGACCGCGGCTTCCGCAACGGGACCTTGAAGATTGGGGGTGACCGG 540			
QY	1026 CGCTGGGAGAGCGTGGGGGCTTACACTTGACATGCGACCAACCTGTGGTGGAGGCGACAGC 1085			
DB	541 CGCTGGGAGAGCGTGGGGGCTTACACTTGACATGCGACCAACCTGTGGTGGAGGCGACAGC 600			
QY	1086 CCGAGTAGAACCTCGGCGGCTGTGCGCTTGGCCCATGTGTGGGAAACAGCATGCGCGAGGCGG 1145			

```

Db      601 CCGAGTGAAGTGGGGTGTCTGGGCTTGGCCCATGTTGGGAACGACAGTGGCGGGGG 660
Qy      1146 CCGCGCCCGGGGCTCTGGACATGCGCGCTCCGCTCGACATGCTGTCGCGAGGGTGAAGGGAC 1205
Db      661 CCGCGCCCGGGGCTCTGGACATGCGCGCTCCGCTCGACATGCTGTCGCGAGGGTGAAGGGAC 720
Qy      1206 GCTGGAGTCTGAGGCGGCGGCTGACAGGTGACGGAGGTGACCGGCCACTCAGGGGCTGGTGA 1265
Db      721 GCTGGAGTCTGAGGCGGCGGCTGACAGGTGACGGAGGTGACCGGCCACTCAGGGGCTGGTGA 780
Qy      1266 CTGGGGTCTCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1325
Db      781 CTGGGGTCTCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Qy      1326 CGAAGATGAGACCTCTATCTACCGGATTTGCTCCAGCTCCAGCCACCACTTCTGCTGAA 1385
Db      841 CGAAGATGAGACCTCTATCTACCGGATTTGCTCCAGCTCCAGCCACCACTTCTGCTGAA 900
Qy      1386 GCACCTGTGCTCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1445
Db      901 GCACCTGTGCTCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
Qy      1446 GCGCTCTGACCTCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1505
Db      961 GCGCTCTGACCTCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
Qy      1506 CTGCGCCCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1565
Db      1021 CTGCGCCCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
Qy      1566 GGGGGGGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1625
Db      1081 GGGGGGGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1140
Qy      1626 GGGGGCGGGAATGAGCGGCTCCCTCAAGCTGAGCGACGCTGCAAGTCCAGCAACAAATGG 1685
Db      1141 GGGGGCGGGAATGAGCGGCTCCCTCAAGCTGAGCGACGCTGCAAGTCCAGCAACAAATGG 1200
Qy      1686 AGGCGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1745
Db      1201 AGGCGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
Qy      1746 CAGCTGTCTCTGTGACCTGGAGATGCGGGTGTCTAGCGTTATGCGAGCGCTGTGGAGG 1805
Db      1261 CAGCTGTCTCTGTGACCTGGAGATGCGGGTGTCTAGCGTTATGCGAGCGCTGTGGAGG 1320
Qy      1806 AGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1865
Db      1321 AGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
Qy      1866 GGTGTGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1908
Db      1381 GTGTGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1423

```

```

RESULT 3
LOCUS   BM552976
DEFINITION
AGNCOURT_6572762 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466979
5', mRNA sequence.
ACCESSION
BM552976
VERSION
BM552976.1 GI:18791300
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1091)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.

```

Email: cgabs-r@mail.nih.gov  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: L109168 row: d column: 20  
 High quality sequence stop: 703.  
 Location/Qualifiers  
 1. 1091  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5466979"  
 /tissue\_type="amelanotic melanoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1lb="NIH MGC 41"  
 /note="Organ: skin, Vector: pOT7, Site 1: XhoI, Site 2:  
 EcoRI, cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 Laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC library."

## ORIGIN

Query Match 48.7%; Score 929.6; DB 4; Length 1091;  
 Best Local Similarity 94.5%; Pred. No. 9.5e-175;  
 Matches 1006; Conservative 0; Mismatches 52; Indels 7; Gaps 4;

```

Qy      817 GAGGGCGAGTTCTCTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 876
Db      1 GAGGGCGAGTTCTCTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
Qy      877 CTGGAAGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 936
Db      61 CTGGAAGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Qy      937 CACTGGGTGGTCTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 996
Db      121 CACTGGGTGGTCTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Qy      997 AACGGGACCTTAAGATTTGGGGTGAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1056
Db      181 AACGGGACCTTAAGATTTGGGGTGAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Qy      1057 GCCACCAACCGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1116
Db      241 GCCACCAACCGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Qy      1117 CATGGTGGAAACAGCACTGCGGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1176
Db      301 CATGGTGGAAACAGCACTGCGGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Qy      1177 GCTCGCACTGCTGCGGAGGGTGAAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1236
Db      361 GCTCGCACTGCTGCGGAGGGTGAAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
Qy      1237 GAGGTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1296
Db      421 GAGGTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
Qy      1297 TGGATGTTCCAAATCCAGTACACAGCGAGGAGATGACCTCATCTACCGGATTTGTC 1356
Db      481 TGGATGTTCCAAATCCAGTACACAGCGAGGAGATGACCTCATCTACCGGATTTGTC 540
Qy      1357 CCGAGCTCCGACGACCACTTCTGCTGAGACACTGTCCTCCGCGGCGGCGGCGGCGGCGGCGGCGG 1416
Db      541 CCGAGCTCCGACGACCACTTCTGCTGAGACACTGTCCTCCGCGGCGGCGGCGGCGGCGGCGGCGG 600

```



QY 1417 TGCTGTGACCTTTGTACCGCGCGTGGGCCCTTGACCTCAAGCCACAGAGCTGCTG 1476  
 Db 601 TGCTGTGACCTTTGTACCGCGCGTGGGCCCTTGACCTCAAGCCACAGAGCTGCTG 660  
 QY 1477 GGCTGTGACCTTTGTACCGCGCGTGGGCCCTTGACCTCAAGCCACAGAGCTGCTG 1536  
 Db 661 GGCTGTGACCTTTGTACCGCGCGTGGGCCCTTGACCTCAAGCCACAGAGCTGCTG 720  
 QY 1537 GTGCTGGGCGGAGCCCTGACCGTGGCGGCGGCGTGGGCCCTTGACCTCAAGCTGCTG 1596  
 Db 721 GTGCTGGGCGGAGCCCTGACCGTGGCGGCGGCGTGGGCCCTTGACCTCAAGCTGCTG 780  
 QY 1597 TTCACTGTGACCTTTGTACCGCGCGTGGGCCCTTGACCTCAAGCTGCTGCTGCTG 1656  
 Db 781 TTCACTGTGACCTTTGTACCGCGCGTGGGCCCTTGACCTCAAGCTGCTGCTGCTGCTG 840  
 QY 1657 CTCAAGCCACAGCTTGTACCGCGCGTGGGCCCTTGACCTCAAGCTGCTGCTGCTGCTG 1716  
 Db 841 CTCAAGCCACAGCTTGTACCGCGCGTGGGCCCTTGACCTCAAGCTGCTGCTGCTGCTG 900  
 QY 1717 CCGCGGAGCCCG 1772  
 Db 901 CCGCGGAGCCCG 960  
 QY 1773 CCGCGTGTACGCTTGTACCGCGCGTGGGCCCTTGACCTCAAGCTGCTGCTGCTGCTGCTG 1829  
 Db 961 CCGCGTGTACGCTTGTACCGCGCGTGGGCCCTTGACCTCAAGCTGCTGCTGCTGCTGCTG 1020  
 QY 1830 TGTGCTGTGACCTTTGTACCGCGCGTGGGCCCTTGACCTCAAGCTGCTGCTGCTGCTG 1874  
 Db 1021 TTTTGTGATGGGCGGCGCTTGTGCGGACGAGGCGGCGCGCGCGCGCGCGCGCGCGCG 1065

## RESULT 4

AL528686 1012 bp mRNA linear EST 24-MAR-2004  
 LOCUS DEFINITION AL528686 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens  
 CDNA clone CS0DD001YC05 5-PRIME, mRNA sequence.  
 AL528686  
 VERSION AL528686.3 GI:45703748  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1012)  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 13, 2001 this sequence version replaced gi:3106536.  
 Contact: Genoscope

Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 6064.f

For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?cs=CS0DD001AB030P1ac=6064.f.

## FEATURES

source  
 1..1012  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DD001YC05"  
 /tissue="NEUROBLASTOMA COT 50-NORMALIZED"  
 /clone\_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 6064.f

## ORIGIN

sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 40.4%; Score 770.8; DB 1; Length 1012;  
 Best Local Similarity 90.8%; Pred. No. 4e-143;  
 Matches 781; Conservative 46; Mismatches 31; Indels 2; Gaps 2;

QY 514 CACACCTCAACCTGACCATTAACCTTATGACGACCTGCGCCAGCGCTTGCGCCAG 573  
 Db 30 CACACCTCAACCTGACCATTAACCTTATGACGACCTGCGCCAGCGCTTGCGCCAG 89  
 QY 574 CTGCGTGAAGTCTCCCGCTGACCTCACTCAACCGCGCTGACGCTGCTGCGAC 633  
 Db 90 CTGCGTGAAGTCTCCCGCTGACCTCACTCAACCGCGCTGACGCTGCTGCGAC 149  
 QY 634 CCGCTTTTCTCTGTGGGCGGTGATGACAGAGCTCTTCGCGCCCGCTGCTGAGCTTT 693  
 Db 150 CCGCTTTTCTCTGTGGGCGGTGATGACAGAGCTCTTCGCGCCCGCTGCTGAGCTTT 209  
 QY 694 AGCGGGAACCCCTGCACTGCAACTGTGAGCTGTGAGCTGTGCGGCGGCGCGCG 753  
 Db 210 AGCGGGAACCCCTGCACTGCAACTGTGAGCTGTGAGCTGTGCGGCGGCGCGCG 269  
 QY 754 GACGACCTGAAAAGTGCAGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 813  
 Db 270 GACGACCTGAAAAGTGCAGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 329  
 QY 814 CCGGAGGCGGAGTTCCTCTGTGAGCGCGCGCTTATGCGCGCACAGCGCGCTTGG 873  
 Db 330 CCGGAGGCGGAGTTCCTCTGTGAGCGCGCGCTTATGCGCGCACAGCGCGCTTGG 389  
 QY 874 GTGCTGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 933  
 Db 390 GTGCTGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 449  
 QY 934 ATGCACTGGGTGCGTCTTGAACGACCGGTTGTTGCAACTCTCCGAGCGCGGCTTTC 993  
 Db 450 ATGCACTGGGTGCGTCTTGAACGACCGGTTGTTGCAACTCTCCGAGCGCGGCTTTC 509  
 QY 994 CCCAAGCGGACCTTAAGATTTGGGCTGACCGCGGCTGCGGAGCGCTGCGGCTTCACTGC 1053  
 Db 510 CCCAAGCGGACCTTAAGATTTGGGCTGACCGCGGCTGCGGAGCGCTGCGGCTTCACTGC 569  
 QY 1054 ATGCGCACCAACCTCTGATGAGGCGCACGCGGTTGTTGCAACTCTCCGAGCGCGGCTTTC 1113  
 Db 570 ATGCGCACCAACCTCTGATGAGGCGCACGCGGTTGTTGCAACTCTCCGAGCGCGGCTTTC 629  
 QY 1114 CCCCATGTGGGAACGACAGTGCAG -GGGAGGCGCGCGCGCGCGCGCGCGCGCGCG 1172  
 Db 630 CCCCATGTGGGAACGACAGTGCAGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCG 689  
 QY 1173 CTCGGTCCGACCTGCGCGGAGGTGAGGAGGAGTGTGAGCTGTGACCGCGGCTGAGGT 1232  
 Db 690 CTCGGTCCGACCTGCGCGGAGGTGAGGAGGAGTGTGAGCTGTGAGCTGTGAGGT 749  
 QY 1233 GACGAGGTGACCGCGACCTCAGGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 1292  
 Db 750 GACGAGGTGACCGCGACCTCAGGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 808  
 QY 1293 AGTGTGATGTTCCAAATCCAGTAAACAAGACGGAAGATGAGACCTTCATCTACCGGAT 1352  
 Db 809 AGTGTGATGTTCCAAATCCAGTAAACAAGACGGAAGATGAGACCTTCATCTACCGGAT 868  
 QY 1353 TGTGCCAGGCTCCAGGACCG 1372  
 Db 869 TGTGCCAGGCTCCAGGACCG 888

## RESULT 5

BUS42726 867 bp mRNA linear EST 13-SEP-2002  
 LOCUS DEFINITION BUS42726 BUS42726 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6574646  
 5', mRNA sequence.

ACCESSION BUS42726  
 VERSION BUS42726.1 GI:22853209  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 867)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: DCTD/DRP  
 DNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
 Plate: LNCM2770 row: e column: 14  
 High quality sequence stop: 708.  
 Location/Qualifiers  
 1..867  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6574646"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH\_MGC\_40"  
 /note="Organ: prostate; Vector: pOTB7; Site:1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."  
 ORIGIN  
 Query Match 40.3%; Score 769; DB 5; Length 867;  
 Best Local Similarity 98.9%; Pred. No. 9e-143;  
 Matches 816; Conservative 0; Mismatches 5; Indels 4; Gaps 4;  
 QY 176 TCATCCGAGCCCTGGGGCCCTGACTTCGCGCAATGAGCGGACTGTGGACCTGAC 235  
 Db 1 TCATCCGAGCCCTGGGGCCCTGACTTCGCGCAATGAGCGGACTGTGGACCTGAC 60  
 QY 236 TGTCTGCAGATGCGATCACCAGATTGGGGCCCGGCTTTGGGGAGCTCGAGAGCTGC 295  
 Db 61 TGTCTGCAGATGCGATCACCAGATTGGGGCCCGGCTTTGGGGAGCTCGAGAGCTGC 120  
 QY 296 GTTCCCTCCACCTTGAAGGCAACAGGCTGTGGAGGCTGGGACCGGAGCCTCCGGGGCC 355  
 Db 121 GTTCCCTCCACCTTGAAGGCAACAGGCTGTGGAGGCTGGGACCGGAGCCTCCGGGGCC 180  
 QY 356 CCGTCATCTGCAGACCTCATCTCTCAGCGGCAACAGCTGCGGATCGCGCGGAG 415  
 Db 181 CCGTCATCTGCAGACCTCATCTCTCAGCGGCAACAGCTGCGGATCGCGCGGAG 240  
 QY 416 CTTTCGACGACTTCTTGAAGAGCTGGAGACCTGACCTGTCTACCAACACTCCGGC 475  
 Db 241 CTTTCGACGACTTCTTGAAGAGCTGGAGACCTGACCTGTCTACCAACACTCCGGC 300  
 QY 476 AGGTGGCCCTGGGCGGGATGGGGCCATGCTGCGCTTGACACCTCAACTGAGCATA 535  
 Db 301 AGGTGGCCCTGGGCGGGATGGGGCCATGCTGCGCTTGACACCTCAACTGAGCATA 360  
 QY 536 ACCATTATTGAGGAGCTGCGGAGGCGCTTGGCCAGCTGGTCACTCTCCGGCTGG 595  
 Db 361 ACCATTATTGAGGAGCTGCGGAGGCGCTTGGCCAGCTGGTCACTCTCCGGCTGG 420

QY 596 ACCTCACCTCCAAACCGCTGGGCAAGCTGGCTCCGACCCGCTTTCCTGTGGGCGTG 655  
 Db 421 ACCTCACCTCCAAACCGCTGGGCAAGCTGGCTCCGACCCGCTTTCCTGTGGGCGTG 480  
 QY 656 ATGACAGAGGCTCTCCGCCCCCTGTGTGTGAGCTTTAAGCGGAAACCCCTGACCTGCA 715  
 Db 481 ATGACAGAGGCTCTCCGCCCCCTGTGTGTGAGCTTTAAGCGGAAACCCCTGACCTGCA 540  
 QY 716 ACTGTAGCTGT 775  
 Db 541 ACTGTAGCTGT 600  
 QY 776 CCGCGCCGCGGCTGGGCGGCGGCGGCTACTTCTGAGGAGTCCGAGGAGGAGTCTCTGTG 835  
 Db 601 CCGCGCCGCGGCTGGGCGGCGGCGGCTACTTCTGAGGAGTCCGAGGAGGAGTCTCTGTG 660  
 QY 836 AGCGGCGGCTGTATTTGCGGCGGCGGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGT 895  
 Db 661 AGCGGCGGCTGTATTTGCGGCGGCGGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGT 720  
 QY 896 CGCTGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 953  
 Db 721 CGCTGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780  
 QY 954 GCAACCGG-TTGTGTGGCACT-CCTCCCGAGCCCGGCGCTTTCCCG 996  
 Db 781 GCAACCGGTTGTGTGGCACTCCTCCCGAGCCCGGCGCTTTCCCG 825  
 RESULT 6  
 LOCUS BQ063824  
 DEFINITION AGENCOURT 6850285 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5925478  
 5', mRNA sequence.  
 ACCESSION BQ063824  
 VERSION BQ063824.1 GI:19891912  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1102)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: Lou Staudt  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
 Plate: LNCM2097 row: d column: 23  
 High quality sequence stop: 689.  
 Location/Qualifiers  
 1..1102  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5925478"  
 /tissue\_type="lymphoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH\_MGC\_99"  
 /note="Organ: lymph; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC

ORIGIN Library."

Query Match 39.1%; Score 745.8; DB 5; Length 1102;  
 Best Local Similarity 94.8%; Pred. No. 3.8e-138;  
 Matches 790; Conservative 0; Mismatches 41; Indels 2; Gaps 2;

QY 786 CTTGCGCGCGCTACTTCTGCGAGTGGCCGAGGGGAGTTCTCTGTGAGCGGCCCT 845  
 1 CTTGCGCGCGCTACTTCTGCGAGTGGCCGAGGGGAGTTCTCTGTGAGCGGCCCT 60

QY 846 CATTGCGCGCGCAACGAGGCGCTCTGCGAGTGGCCGAGGGGAGTTCTCTGTGAGCGGCCCT 905  
 61 CATTGCGCGCGCAACGAGGCGCTCTGCGAGTGGCCGAGGGGAGTTCTCTGTGAGCGGCCCT 120

QY 906 CCGGCGCGCTGAGTGGCCGCGCTGAGTGGCCGAGGGGAGTTCTCTGTGAGCGGCCCT 965  
 121 CCGGCGCGCTGAGTGGCCGCGCTGAGTGGCCGAGGGGAGTTCTCTGTGAGCGGCCCT 180

QY 966 TGGCAACTCTCTCCGAGCGCGGCTTTCCCAACGAGGAGCTTGAAGATTGGGCTGACCGG 1025  
 181 TGGCAACTCTCTCCGAGCGCGGCTTTCCCAACGAGGAGCTTGAAGATTGGGCTGACCGG 240

QY 1026 CCGTGGGAGAGCGTGGGCGCTGACCTGCAATGCGCAACCACTGCTGAGGCGCAACG 1085  
 241 CCGTGGGAGAGCGTGGGCGCTGACCTGCAATGCGCAACCACTGCTGAGGCGCAACG 300

QY 1086 CCGAGTAGAAGTGGGCGCTGAGTGGCCGAGGGGAGTTCTCTGTGAGCGGCCCT 1145  
 301 CCGAGTAGAAGTGGGCGCTGAGTGGCCGAGGGGAGTTCTCTGTGAGCGGCCCT 360

QY 1146 CCGCGCGCGCGCTCTGCAATGCGCGCTCTGCACTGCTGCGAGGCGTGGGAGG 1205  
 361 CCGCGCGCGCGCTCTGCAATGCGCGCTCTGCACTGCTGCGAGGCGTGGGAGG 420

QY 1206 GCTGAGTGGAGCGCGCTGAGTGGCCGAGGGGAGTTCTCTGTGAGCGGCCCT 1265  
 421 GCTGAGTGGAGCGCGCTGAGTGGCCGAGGGGAGTTCTCTGTGAGCGGCCCT 480

QY 1266 CTTGCG 1325  
 481 CTTGCG 540

QY 1326 CCGAAGTAGAAGCTCTATCTACCGAGTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1385  
 541 CCGAAGTAGAAGCTCTATCTACCGAGTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600

QY 1386 GCAACCTGTCG 1445  
 601 GCAACCTGTCG 660

QY 1446 GCGCTTGAAGTGGAGCG 1504  
 661 GCGCTTGAAGTGGAGCG 720

QY 1505 CCTGCGCGCGCTGTCG 1563  
 721 CCTGCGCGCGCTGTCG 780

QY 1564 GTGGGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1616  
 781 CCGCTGCGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 833

RESULT 7  
 BUS38343 953 bp mRNA linear EST 13-SEP-2002  
 LOCUS AGNCOUNT\_10181110 NIH\_MGC\_107 Homo sapiens cDNA clone  
 DEFINITION IMAGE:5568883 5', mRNA sequence.  
 ACCESSION BUS38343  
 VERSION BUS38343.1 GI:22848784  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 953)  
 AUTHORS NIH-MGC  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaab@nhi.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: L1CM2755 row: e column: 11  
 High quality sequence stop: 465.  
 Location/Qualifiers  
 1..953  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5568883"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 107"  
 /note="Organ: breast; Vector: pOTB7; Site: 1: EcoRI;  
 Site: 2: XhoI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(5). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 38.6%; Score 737.4; DB 5; Length 953;  
 Best Local Similarity 92.9%; Pred. No. 1.8e-136;  
 Matches 796; Conservative 0; Mismatches 56; Indels 5; Gaps 2;

QY 99 CTTGCG 158  
 1 CTTGCG 60

QY 159 GCGGCTGCGTGAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 218  
 61 GCGGCTGCGTGAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 120

QY 219 ACTGTGAGCTGAGACTGTCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 278  
 121 ACTGTGAGCTGAGACTGTCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 180

QY 279 GGAACCTGAGAGCTGCGTCTCTCTCACTTGAAGCGCAACAGGCTGGAGCTGGAGC 338  
 181 GGAACCTGAGAGCTGCGTCTCTCTCACTTGAAGCGCAACAGGCTGGAGCTGGAGC 240

QY 339 CGGAGCGCTCGGCG 398  
 241 CGGAGCGCTCGGCG 300

QY 399 CCGCATGCG 458  
 301 CCGCATGCG 360

QY 459 CTACAAACAATCTTCGCGAGTGGCTGGGCGCGCATGCGCGCATGCGCGCATGCGCGCATGCGCGCATGCGCG 518  
 361 CTACAAACAATCTTCGCGAGTGGCTGGGCGCGCATGCGCGCATGCGCGCATGCGCGCATGCGCGCATGCGCG 420

QY 519 CCTCAACCTGAGCACTTAATTTAGAGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 578  
 421 CCTCAACCTGAGCACTTAATTTAGAGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480

QY 579 TCAAGCTTCGCGCTGAGACCTCACTCCACACCGCTGGCCAGCTGGCTCCGACCCGCT 638  
 Db 481 TCGAGCTTCGCGCTGAGACCTCACTCCACACCGCTGGCCAGCTGGCTCCGACCCGCT 540  
 QY 639 TTTCTCTCGTGGGCGTGTATGACAGAGGCTCTCCGACCCCTGGTGTGTAGCTTTAGCG 638  
 Db 541 TTTCTCTCGTGGGCGTGTATGACAGAGGCTCTCCGACCCCTGGTGTGTAGCTTTAGCG 600  
 QY 699 GAACCCCTGACCTGACACTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 758  
 Db 601 GAACCCCTGACCTGACACTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660  
 QY 759 CCGAAGAGCTGGCGCTCCCGCGCGCGCTGGCGCGCTGGCGCGCTGGCGCGCTGGCGCG 818  
 Db 661 CCGAAGAGCTGGCGCTCCCGCGCGCGCTGGCGCGCTGGCGCGCTGGCGCGCTGGCGCG 720  
 QY 819 GGGGAGATTCCTCTGTGTAGCGCGCGCTCATTTGCCCGACAGCGAGCGC--TCTGGGTG 876  
 Db 721 GGGGAGATTCCTCTGTGTAGCGCGCGCTCATTTGCCCGACAGCGAGCGC--TCTGGGTG 780  
 QY 877 CTGGAGAGCGCAGCGCGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 933  
 Db 781 TGAGAAAGCCATTGGGCGCCCTTTCCGCGCGCGCGCGCTGGGTGGAATCCGCCCTAAC 840  
 QY 934 ATGACCTGGGTCGCTCC 950  
 Db 841 ATGACCTGGGTCGCTCC 857

RESULT 8  
 AL563226/c 1085 bp mRNA linear EST 05-APR-2004  
 LOCUS AL563226 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens  
 DEFINITION CDNA clone CS0D001YC05 3-PRIME, mRNA sequence.  
 ACCESSION AL563226  
 VERSION AL563226.3 GI:46230280  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1085)  
 Li, W. B., Gruber, C., Jesse, J., and Polayres, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 15, 2001 this sequence version replaced gi:31287230.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 6064.f

For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?cs=CS0D001AB03NP1&c=6064.f.

FEATURES  
 source Location/Qualifiers  
 1..1085

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D001YC05"  
 /issue\_type="NEUROBLASTOMA COT 50-NORMALIZED"  
 /clone\_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Query Match 37.5%; Score 714.6; DB 1; Length 1085;

Best Local Similarity 93.7%; Pred. No. 6 2e-132;  
 Matches 749; Conservative 21; Mismatches 26; Indels 3; Gaps 3;

QY 1110 CTTGCCCATGATGAGGAGACAGATGCGCAGGAGGAGCGCGCCCGGCGCTTGACATCGC 1169  
 Db 1059 CTTGGCCACTGATGAGACAGAGTGTGCGA--GGGAGCGCGCCCGGCGCTTGAGAAATCGC 1001  
 QY 1170 CGCTCCGCTGACAGCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1229  
 Db 1000 MGATCAGGTGAGCATGCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 941  
 QY 1230 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1289  
 Db 940 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 882  
 QY 1290 CCGAGTGTGATGTTCCAAATCCAGTACAAAGACAGGAGGAGGAGGAGGAGGAGGAGGAG 1349  
 Db 881 CCGAGTGTGATGTTCCAAATCCAGTACAAAGACAGGAGGAGGAGGAGGAGGAGGAGGAG 822  
 QY 1350 GATTGTCCAGGCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1409  
 Db 821 GATTGTCCAGGCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 762  
 QY 1410 TGAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1469  
 Db 761 TGAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702  
 QY 1470 GTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1529  
 Db 701 GTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 643  
 QY 1530 GGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1589  
 Db 642 GGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 583  
 QY 1590 ACTGATCTTCACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1649  
 Db 582 ACTGATCTTCACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 523  
 QY 1650 CTTCAAGCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1709  
 Db 522 CTTCAAGCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 463  
 QY 1710 CCAACCGCGCGGAGCG 1769  
 Db 462 CCAACCGCGCGGAGCG 403  
 QY 1770 TGCAGGAGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1829  
 Db 402 TGCAGGAGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 343  
 QY 1830 TGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1889  
 Db 342 TGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 283  
 QY 1890 GGAAGAGAGTGTGTGTGA 1908  
 Db 282 GGAAGAGAGTGTGTGTGA 264

RESULT 9  
 BM915392 996 bp mRNA linear EST 12-MAR-2002  
 LOCUS BM915392  
 DEFINITION AGENCOURT\_6701786 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5481569  
 5', mRNA sequence.  
 ACCESSION BM915392  
 VERSION BM915392.1 GI:19365771  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 996)



Db	1	AAGAAGCTGCGGGTGTCTGGCCCTTGTGCCCCATGAGTGGAGAACGACAGTGGCCGAGGGGGGCGCGCC	60
QY			
Db	1151	CCGGGGCCCTCGGACATATGCGGCCCTTCCTGCTCGGCACTGCTGCTCCGAGAGGTGAGGGGGACGCTGG	1210
QY			
Db	61	CCGGAGCCCTTGGACATATGCGGCCCTTCCTGCTCGGCACTGCTGCTCGAGAGGTGAGGGGGACGCTGG	120
QY	1211	AGTGTGAGCCAGCCGCTGAGGTGACGAGAGTGAACGAGCGGACCTCAAGGGCTGTGTGACTGGG	1270
Db	121	AGTGTGAGCCAGCCGCTGAGGTGACGAGAGTGAACGAGCGGACCTCAAGGGCTGTGTGACTGGG	180
QY			
Db	1271	GTCCCGGGCGGGCCAGCCGACCACTGATGTGTGATGTTTCCAAATCCAGTATCAACAGCCGAG	1330
QY			
Db	181	GTCCCGGGCGGGCCAGCCGACCACTGATGTGTGATGTTTCCAAATCCAGTATCAACAGCCGAG	240
QY	1331	ATGAGACCCCTATCTACCGGATTGTGCCAGGCTCCAGGACCACTCTCTGTGTAAGCACG	1390
Db	241	ATGAGACCCCTATCTACCGGATTGTGCCAGGCTCCAGGACCACTCTCTGTGTAAGCACG	300
QY			
Db	1391	TGTGTCCCGGGCGCTGACTATGACTCTGTGCTGTGAGCTTGTACCGGCGGCTGTGGCCCT	1450
QY			
Db	301	TGTGTCCCGGGCGCTGACTATGACTCTGTGCTGTGAGCTTGTACCGGCGGCTGTGGCCCT	360
QY			
Db	1451	CTGACTACAGGACCAACGAGCTGCTGGGCTGTGCCATTCTTCCACGCTGGCGGCTGTGG	1510
QY			
Db	361	CTGACTACAGGACCAACGAGCTGCTGGGCTGTGCCATTCTTCCACGCTGGCGGCTGTGG	420
QY	1511	CCCTGTGCGACGCGCTGTGAGGCCACGCTGTGAGGGGAGCCCTGACCCGTGAGCCGTGGGG	1570
Db	421	CCCTGTGCGACGCGCTGTGAGGCCACGCTGTGAGGGGAGCCCTGACCCGTGAGCCGTGGGG	480
QY			
Db	1571	GTGTGCTGTGTGTGCTGCTTACTGTGTCTTCACTGTGTGCTTGTGTTC--GGGGCGGGGG	16229
QY			
Db	481	GTGTGCTGTGTGTGCTGCTTACTGTGTCTTCACTGTGTGCTTGTGTTC--GGGGCGGGGG	540
QY	1630	GCCGGAATAGGCGCGCTCCCTCCCTCAAGCTCAGCCACGCTCCAGTCCCAACCAATGAGAGC	1689
Db	541	GCCGGAATAGGCGCGCTCCCTCCCTCAAGCTCAGCCACGCTCCAGTCCCAACCAATGAGAGG	600
QY	1690	CCCAAGCCCCACACCCAGGCCACCC--GCCGCGAGGCCCTCCGCGCGGCGGACGCTGAG	1748
Db	601	CCCAAGCCCCACACCCAGGCCACCC--GCCGCGAGGCCCTCCGCGCGGCGGACGCTGAG	660
QY	1749	CTGCTCTCTGTGACCTGTGGAGATGAC--GGGTGTGCTACGTTATGTCCAGAGCGGCTGTGGAGAG	1807
Db	661	CTGCTCTCTGTGACCTGTGGAGATGAC--GGGTGTGCTACGTTATGTCCAGAGCGGCTGTGGAGAG	720
QY	1808	CTTGGAGCCGACGAGGCCACTTGTGTGATGAGGGGGGCTGTGGGGCAGAGTGCCTGGGGGGG	1867
Db	721	CTTGGAGCCGACGAGGCCACTTGTGTGATGAGGGGGGCTGTGGGGCAGAGTGCCTGGGGGGG	780
QY			
Db	1868	TA 1869	
QY			
Db	781	GA 782	
RESULT 11			
LOCUS	BM046297	708 bp	mRNA linear EST 07-NOV-2001
DEFINITION	603626095E1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:545630 5'		
ACCESSION	BM046297		
VERSION	BM046297.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 708)		
TITLE	NIH-MGC http://mgi.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs-1@mail.nih.gov		

ORIGIN	Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
QY	174	CTTATCCAGGCGCTCGGGGCCCTTGACTTCCGCACATGACGGAGCTGTGTGACACTGAC	233				
Db	2	CTTATCCAGGCGCTCGGGGCCCTTGACTTCCGCACATGACGGAGCTGTGTGACACTGAC	61				
QY	234	ACTGCTCGCAGTGGCATCACCCGCGATTGGGGACCGCGCCTTTGGGGACCTCGAGAGCCT	293				
Db	62	ACTGCTCGCAGTGGCATCACCCGCGATTGGGGACCGCGCCTTTGGGGACCTCGAGAGCCT	121				
QY	294	GCGTTCCCTCCACCTTGAAGGCAACGAGCTGGTGGAGCTGGGCAACCGGAGCTTCGCGG	353				
Db	122	GCGTTCCCTCCACCTTGAAGGCAACGAGCTGGTGGAGCTGGGCAACCGGAGCTTCGCGG	181				
QY	354	CCCCGTCAATGAGCAACGCTCATCTGAGGGGGAACAAGCTGGGCGGCGATGCGCGCGG	413				
Db	182	CCCCGTCAATGAGCAACGCTCATCTGAGGGGGAACAAGCTGGGCGGCGATGCGCGCGG	241				
QY	414	AGCCTTCAGCACATCTCTAGAGAGCGCTGAGAGCACTGTGACCTGTCTTCAACACACCTCG	473				
Db	242	AGCCTTCAGCACATCTCTAGAGAGCGCTGAGAGCACTGTGACCTGTCTTCAACACACCTCG	301				
QY	474	GAGGTGCGCTGGGCGGCGCATGTGGGCGCATGTCTGCGCTTGACAACCTTCACCTGGACCA	533				
Db	302	GAGGTGCGCTGGGCGGCGCATGTGGGCGCATGTCTGCGCTTGACAACCTTCACCTGGACCA	361				
QY	534	TAACTTATTAAGCAGCAGCGCCCGCAGGGGCTTCCGCGAGCTCGGTACGCTTCCGCGCT	593				
Db	362	TAACTTATTAAGCAGCAGCGCCCGCAGGGGCTTCCGCGAGCTCGGTACGCTTCCGCGCT	421				
QY	594	GAGCTCACTTCAACCGCGCTGGGCGACGCTGGCTCCGGAACCGGCTTTTCTCTGTGGGGG	653				
Db	422	GAGCTCACTTCAACCGCGCTGGGCGACGCTGGCTCCGGAACCGGCTTTTCTCTGTGGGGG	481				
QY	654	TGATGCAAGGCGCTTCCGCGCGCGCTGGTGTGAGCTTTAGCGGGAACCCCTGACACTG	713				
Db	482	TGATGCAAGGCGCTTCCGCGCGCGCTGGTGTGAGCTTTAGCGGGAACCCCTGACACTG	541				
QY	714	CAACTGTAGCTGTGTGGCTGCGGCGAGCTGGCGGCGCGACGACTTGGAAAGCTGCGC	773				
Db	542	CAACTGTAGCTGTGTGGCTGCGGCGAGCTGGCGGCGCGACGACTTGGAAAGCTGCGC	601				
QY	774	CTCCCGCGCGGCTGGCGCGGCTACTTCTGGGCAATGCCCGAGGGGGAATTTCTCTCG	833				

```

|||||
Db      602 CTCCTCCGACCGGCTGCGCGCGCTACTTCTGGGCGAGTGGCGAGGCGAGTTCTGCTG 661
QY      834 TGAAGCGCCCTGATTTGCGCGCGACAGCAGCGCGCTTGGGTCTG 880
Db      662 TGAAGCTGCGCTCATTTGCGCGACAGCAGCGCGCTTGGGTCTG 708

RESULT 12
LOCUS   B0062401
DEFINITION B0062401 1086 bp mRNA linear EST 02-APR-2002
5' mRNA sequence.
ACCESSION B0062401
VERSION   B0062401.1 GI:19889199
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1086)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT    Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM0293 row: b column: 03
High quality sequence stop: 475.

FEATURES
Source
Location/Qualifiers
1..1086
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5923874"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 99"
/note="Organ: lymph. Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH-MGC
Library."

ORIGIN
Query Match 35.7%; Score 680.6; DB 5; Length 1086;
Best Local Similarity 95.1%; Pred. No. 3.6e-125;
Matches 758; Conservative 0; Mismatches 29; Indels 10; Gaps 5;

QY 790 GCGGCGCTACTTCTGGGAGTGGCGAGGCGAGTTCTCTGTGAGCCGCTCAT 849
Db 1 GCGGCGCTACTTCTGGGAGTGGCGAGGCGAGTTCTCTGTGAGCCGCTCAT 60
QY 850 GCGGCGCACAGGAGGCTCTGGGTGTGAGGCGAGGCGAGCTGGCGGCGCG 909
Db 61 GCGGCGCACAGGAGGCTCTGGGTGTGAGGCGAGGCGAGCTGGCGGCGCG 120
QY 910 GCGCTGGGTGACCGCGCGCTTACCATGCACTGGGTGTGAGGCGAGTGTG 969
Db 121 GCGCTGGGTGACCGCGCGCTTACCATGCACTGGGTGTGAGGCGAGTGTG 180
QY 970 AATCTCTCCGAGCCCGGCTTCCCAAGGAGCTTAGAGTTGGGGTGAACCGCGCT 1029

```

```

Db      181 AACTCTCCCGAGCCCGGCTTCCCAAGGAGCTTAGAGTTGGGTGACCGCGCT 240
QY      1030 GGGGAGGCTGGGGGCTACACCTGACATCGCCACCAACCTGCTGTGAGGCGACAGCCGA 1089
Db      241 GGGGAGGCTGGGGGCTACACCTGACATCGCCACCAACCTGCTGTGAGGCGACAGCCGA 300
QY      1090 GTAGAACTGGGGTGTGAGCTTGGCCCAATGGTGGGAGACAGAGTGGCGGAGGGGCGG 1149
Db      301 GTAGAACTGGGGTGTGAGCTTGGCCCAATGGTGGGAGACAGAGTGGCGGAGGGGCGG 360
QY      1150 CCCGGGCGCTCGACATCGCCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTG 1209
Db      361 CCCGGGCGCTCGACATCGCCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTG 420
QY      1210 GAGTCTGAGCCAGCGCTGACAGTGAAGAGAGTGAACCGCACTCAGGAGCTGTGAGCTGG 1269
Db      421 GAGTCTGAGCCAGCGCTGACAGTGAAGAGAGTGAACCGCACTCAGGAGCTGTGAGCTGG 480
QY      1270 GGTCCCGGCGGCGGACCGACCGACCGAGTGGATGTCCTCAATCCAGTACACAGCGGAA 1329
Db      481 GGTCCCGGCGGCGGACCGACCGACCGAGTGGATGTCCTCAATCCAGTACACAGCGGAA 540
QY      1330 GATGAGACCTCTATCTACCGGATTTGCCAGCCTCCAGCCACCACTTCTGTGAAGCAG 1389
Db      541 GATGAGACCTCTATCTACCGGATTTGCCAGCCTCCAGCCACCACTTCTGTGAAGCAG 600
QY      1390 CTGCTCCCGGCGGCTGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1449
Db      601 CTGCTCCCGGCGGCTGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY      1450 TCTGACCTGAC--GGCCACAGGCTGTGGG--CTGTGCCATTCTC--CAGCTGCGGG-- 1504
Db      661 TCTGACCTGACAGGCGACAGGAGTGTGGGCTGTGCCATTCTCAGCTGCGGGG 720
QY      1505 CCTGCGCCTGTGTCACGCGCC---TGAAGGCCACGTGCTGGGCGGACCTTGACCT 1559
Db      721 CTTGCGCCTGTGTCACGCGCCCTGAAGTCCCACTTGTGCGGCGGCGACCTTGATCG 780
QY      1560 GGCCGCTGGGGGCTGTGC 1576
Db      781 GGCCGCTGGGGGCTGTGC 797

RESULT 13
LOCUS   BM784343
DEFINITION K-EST0062439 S7SNU719 Homo sapiens cDNA clone S7SNU719-34-C06 5',
mRNA sequence.
ACCESSION BM784343
VERSION   BM784343.1 GI:19132575
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 703)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
CONTACT   Kim YS
GENOME RESEARCH CENTER
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.krrib.re.kr
Plate: 34 row: C column: 06
High quality sequence stop: 703.

FEATURES
Source
Location/Qualifiers
1..703
/organism="Homo sapiens"

```



ORIGIN

Query Match	35.6%;	Score 679;	DB 4;	Length 703;
Best Local Similarity	99.7%;	Pred. No. 7.4e-125;		
Matches 701; Conservative	0;	Mismatches 0;	Indels 2;	Gaps 2

QY	372	CCTCATCTCAGCGGAAACAGCTGAGGCGCATCGCGCGGAGGCTTCAACGATTCT	431
Db	1	CCTCATCTCAGCGGAAACAGCTGAGGCGCATCGCGCGGAGGCTTCAACGATTCT	60
QY	432	AGAGAGCTTGAGGAGCACTGGAAGCTGTCTTAACAACACTTCGCGCAGTGTCCCTGGGCGG	491
Db	61	AGAGAGCTTGAGGAGCACTGGAAGCTGTCTTAACAACAACCTCCGGCAGTGTCCCTGGGCGG	120
QY	492	CATGAGCGGCATGCGCTGCGCAACCCCTCAACCTGGACCATTAACCTTATTGAGCACT	551
Db	121	CATGAGCGGCATGCGCTGCGCAACCCCTCAACCTGGACCATTAACCTTATTGAGCACT	180
QY	552	GCCCCAAGGCGCTTGTGCGCAGCTTGAGCTCTCCGCTGAGACTCACTCCACCG	611
Db	181	GCCCCAAGGCGCTTGTGCGCAGCTTGAGCTCTCCGCTGAGACTCACTCCACCG	240
QY	612	CGTGGCAAGCTGGGCTCCGGAACCGGCTTTCTGTGTGGGGGTGATGAGAAGGCGCTTCC	671
Db	241	CGTGGCAAGCTGGGCTCCGGAACCGGCTTTCTGTGTGGGGGTGATGAGAAGGCGCTTCC	300
QY	672	CGCCCCCTGTGTGAGCTTTAAGGGAAACCCCTGACCTGACCACTGTGAGCTGTGTG	731
Db	301	CGCCCCCTGTGTGAGCTTTAAGGGAAACCCCTGACCTGACCACTGTGAGCTGTGTG	360
QY	732	GCTCGGGGAGCTGGGCGGGCCGGACGACTGTGAAACGTGGCGCTCCCGCCCGGCTGGC	791
Db	361	GCTCGGGGAGCTGGGCGGGCCGGACGACTGTGAAACGTGGCGCTCCCGCCCGGCTGGC	420
QY	792	CGGCGCTACTTCTGGGAGTGCCTCGAAGGCGAATTCTCTGTGAGCGGCCCTCATTTGC	851
Db	421	CGGCGCTACTTCTGGGAGTGCCTCGAAGGCGAATTCTCTGTGAGCGGCCCTCATTTGC	480
QY	852	CCGCGACAACGAGGCGCTGTGGGTGTGAAAGCCAGGGGGGCAAGCTGTGGGTGCGGGCC	911
Db	481	CCGCGACAACGAGGCGCTGTGGGTGTGAAAGCCAGGGGGGCAAGCTGTGGGTGCGGGCC	540
QY	912	CTTGGGTGACCCCGCG-CTTACCAATGCACTGGGTGCTCTTGACG-ACCGATTGTGTGGC	969
Db	541	CTTGGGTGACCCCGCGCTTACCAATGCACTGGGTGCTCTTGAGCAACCGATTGTGTGGC	600
QY	970	AATCTCTCCGAGCCCGGAGCTTTCCCAACGGAGACTTAGAATTGGGGTGAACGGGCGCT	102

Db 601 AATCTCTCCCGAGCCCGGGCTTTTCCCAACGGGACCTTAGAGATTTGGGGTGAGCCGGCGT 660

Dy 1030 GGGGACGCTGGGGGCTACACGTGATGGCCACCAACCCCTGCTG 1072

Db 661 GGGGACGCTGGGGGCTACACGTGATGGCCACCAACCCCTGCTG 703

RESULT	14
BM722122	
LOCUS	
DEFINITION	674 bp mRNA linear EST 01-MAR-2002
	UT-E-B00-ahy-a-11-0-UI.r1 UT-E-B00 Homo sapiens cDNA clone
	UT-E-B00-ahy-a-11-0-UI.5, mRNA sequence.

ACCESSION	BM722122
VERSION	BM722122.1
KEYWORDS	EST.
SOURCE	Homo sapiens (human)

ORGANISM	REFERENCE
Homo sapiens	1 (bases 1 to 674)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi	
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	

**AUTHORS**      **TITLE**  
 Ronaldo, M. F., Lennon, G., and Soares, M. B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
**JOURNAL**      **Genome Res.** 6 (9), 791-806 (1996)

MEDLINE 97044477  
PUBMED 8889548  
COMMENT Contact: Soares, MB

Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250

Fax: 319 335 9565  
Email: bento-soares@iowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics ([www.resgen.com](http://www.resgen.com))

```
Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source            1..674
```

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="UI-E-B00-ahy-a-11-0-UI"
```

```

/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_id="UI-E-E00"

```

/note=Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site\_1: BcoR I; Site\_2: Not I; UI-E-E00 is a cDNA library containing the following tissue(s): feline eye. The library was constructed

according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)<sub>18</sub> tail. The

sequence tag for this library is GCGGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI). "

Query Match	35.2%	Score 670.8	DB 4	Length 674
Best Local Similarity	99.7%	Pred. No. 3.2e-123		
Matches 672; Conservative	0	Mismatches 2	Indels 0	Gaps 0

289 AGCCTGCGTTCCCTCCACCTTGACGGCAACAGAGCTGTGAGCTGGGCACCCGGAGCCTC 348

```

Db      1 AGCTGCGTTCCCTCCACCTTGAACGGCAACAGGCTGTGAGAGCTGGGACACCGGAGCCTC 60
QY      349 CGGAGCCCGCTCAATCTGACGACCTATCTCTAGCGGCAACAGCTGGGCGCATGCG 408
Db      61 CGGGGCCCCCTCAATCTGACGACCTATCTCTAGCGGCAACAGCTGGGCGCATGCG 120
QY      409 CGGAGCCCTTGAACGACCTTCTAGAGAGCTTGAAGAGCTTGAAGAGCTTGTCTTCAACAC 468
Db      121 CGGAGAGCTTGAACGACCTTCTAGAGAGCTTGAAGAGCTTGTCTTCAACAC 180
QY      469 CTCGGGAGAGTGGCTGGGCGGCGATGCGGCGCATGCTGCTGACACCTTCAACCTG 528
Db      181 CTCGGGAGAGTGGCTGGGCGGCGATGCGGCGCATGCTGCTGACACCTTCAACCTG 240
QY      529 GACCATTAACCTTATTAAGACGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 588
Db      241 GACCATTAACCTTATTAAGACGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY      589 CGCTGAGACCTCACTTCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 648
Db      301 CGCTGAGACCTCACTTCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY      649 GGGCGGAGTGCAGAGGCGCTCTCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 708
Db      361 GGGCGGAGTGCAGAGGCGCTCTCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY      709 CACTGCAACTGTGTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 768
Db      421 CACTGCAACTGTGTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
QY      769 TGGCGCTCCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 828
Db      481 TGGCGCTCCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
QY      829 TCGTGTGAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 888
Db      541 TCGTGTGAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
QY      889 CGGAGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 948
Db      601 CGGAGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
QY      949 CCGTGAACGACCGGTT 962
Db      661 CCGTGAACGACCGGTT 674

RESULT 15
CD358944          915 bp      mRNA      linear      EST 29-MAY-2003
LOCUS             CD358944
DEFINITION       AGENCOURT 14276726 NIH MGC 180 Homo sapiens cDNA clone
ACCESSION        CD358944
VERSION          CD358944.1 GI:31130355
KEYWORDS         EST.
SOURCE            Homo sapiens (human)
ORGANISM         Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 915)
NIH-MGC http://mgc.nci.nih.gov/
Unpublished (1999)
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

```

```

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30389020"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
/clone_lib="NIH MGC 180"
/note="Organ: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI;
Site 2: EcoRV (destroyed); Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.68 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match      34.4%; Score 655.6; DB 6; Length 915;
Best Local Similarity 98.2%; Pred. No. 3.4e-120;
Matches 695; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

```

```

616 GCCACGCTGGCTCCGACCCGCTTCTCGTGGGCGTGAATGACAGGCGCTTCCGCG 675
Db      5 GGGCGCGTGGCTTCGACCCGCTTCTCGTGGGCGTGAATGACAGGCGCTTCCGCG 64
QY      676 CCCCTGTGTCTAGCTTTAGCGGAAACCCCTGCACTGCAACTGTGAGCTGTGTGTGT 735
Db      65 CCCCTGTGTCTAGCTTTAGCGGAAACCCCTGCACTGCAACTGTGAGCTGTGTGTGT 124
QY      736 CGGCGGCTGGCGCGCGCGGAGACGACCTGGAAGCTGCGCTCCCGCGCGGCGGCGGCGG 795
Db      125 CGGCGGCTGGCGCGCGCGGAGACGACCTGGAAGCTGCGCTCCCGCGCGGCGGCGG 184
QY      796 CGCTACTTCTGGGAGATGCCGAGGCGGATCTCTGAGAGCGGCGGCGGCGGCGGCGG 855
Db      185 CGCTACTTCTGGGAGATGCCGAGGCGGATCTCTGAGAGCGGCGGCGGCGGCGGCGG 244
QY      856 CACAGCAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 915
Db      245 CACAGCAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 304
QY      916 GGTGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 975
Db      305 GGTGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 364
QY      976 TCCCGAGCCCGGCGGCTTCCCAACGAGACCTTGAAGATGAGGATGAGCGGCGTGGGAG 1035
Db      365 TCCCGAGCCCGGCGGCTTCCCAACGAGACCTTGAAGATGAGGATGAGCGGCGTGGGAG 424
QY      1036 GCTGGGAGGCTTACACCTGTGACGACCAACCTGTGTGTGTGTGTGTGTGTGTGTGT 1095
Db      425 GCTGGGAGGCTTACACCTGTGACGACCAACCTGTGTGTGTGTGTGTGTGTGTGTGT 484
QY      1096 CTGCGGAGTGTGCGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1155
Db      485 CTGCGGAGTGTGCGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 544
QY      1156 CCTCGGACATGCGCGGCTTCCGCTGCACTGTGTGCGA-GGGTGAAGGAGCGCTGAGATC 1214
Db      545 CCTCGGACATGCGCGGCTTCCGCTGCACTGTGTGCGAAGGGGTGAAGGAGCGCTGAGATC 604
QY      1215 TAGAGCAGCGGTGACAGTGAAGAGTGAACCGGCGGCGGCGGCGGCGGCGGCGGCGG 1274
Db      605 TAGAGCAGCGGTGACAGTGAAGAGTGAACCGGCGGCGGCGGCGGCGGCGGCGGCGG 664
QY      1275 C-GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1319
Db      665 CGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 712

```

Sat Nov 20 16:15:42 2004

us-10-071-879-9.rst

Page 14

Search completed: November 20, 2004, 14:47:35  
Job time : 8936 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 04:30:49 ; Search time 1327 Seconds  
(without alignments)  
7547.773 Million cell updates/sec

Title: US-10-071-879-9

Perfect score: 1908

Sequence: 1 atggcccgccgcgtcctcgtc.....tggaagagagtgtrgtga 1908

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4134886 segs, 2624710521 residues 8269772

Total number of hits satisfying chosen parameters:

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

1: N\_Geneseq\_23Sep04:\*  
2: geneseqn1980s:\*  
3: geneseqn1990s:\*  
4: geneseqn2000s:\*  
5: geneseqn2001as:\*  
6: geneseqn2002as:\*  
7: geneseqn2002ds:\*  
8: geneseqn2003as:\*  
9: geneseqn2003ds:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1906.4	99.9	2369	6 AAD38695	Aad38695 Human LP2
2	1906.4	99.9	2569	10 AAK99609	AAK99609 MDDT rela
3	1904.8	99.8	2364	10 ADE07450	ADe07450 Novel cod
4	1897	99.4	3201	6 ABV99420	ABv99420 Human NOV
5	1789.6	93.8	2885	10 ADI21820	ADI21820 Novel hum
6	1787.6	93.7	2563	6 ABV99422	ABv99422 Human NOV
7	1629	85.4	1653	6 AAD28124	Aad28124 Human leu
8	1584.4	83.0	2451	6 ABV99421	ABv99421 Human NOV
9	1563.4	81.9	1872	4 AAF76853	Aaf76853 Human sec
10	1379.4	72.3	1582	5 AAS69014	Aas69014 DNA encod
11	1101.8	57.7	1653	6 AAD28125	Aad28125 Human leu
12	624.8	32.7	1887	4 AAI66986	Aai66986 Human LRR
13	624.8	32.7	1887	4 AAH78215	Aah78215 Nucleotid
14	624.8	32.7	1887	4 AAH25568	Aah25568 Nucleotid
15	624.8	32.7	1887	12 ADH71653	Adh71653 Human gen
16	624.8	32.7	1887	12 ADH71651	Adh71651 Human gen
17	624.8	32.7	2549	10 ADF69163	Adf69163 Human MP5
18	624.8	32.7	2637	4 AAI66985	Aai66985 Human leu
19	624.8	32.7	3124	6 ABZ11543	Abz11543 Human pol
20	624.8	32.7	3245	12 ADM44061	Adm44061 Novel hum
21	621	32.5	3144	3 AAA60605	Aaa60605 Human thio

22	621	32.5	3144	3 AAA40083	Aaa40083 Human bra
23	621	32.5	3144	10 ADF69162	Adf69162 Human MP5
24	619.4	32.5	2818	5 AAI58215	Aai58215 Human pol
25	619.4	32.5	2818	5 AAS44992	Aas44992 cDNA enco
26	619.4	32.5	2818	5 ADQ98422	Adq98422 DNA enco
27	619.4	32.5	2818	5 ADB48182	Adb48182 Novel hum
28	617	32.3	2316	4 AAH78204	Aah78204 Nucleotid
29	617	32.3	2869	10 AAD54310	Aad54310 Human SEC
30	617	32.3	3386	6 ABT08315	Abt08315 Human NOV
31	617	32.3	3568	6 ABT08317	Abt08317 Human NOV
32	617	32.3	4042	10 ADI21778	ADI21778 Novel hum
33	612	32.1	2855	6 ABK70010	Abk70010 cDNA enco
34	612	32.1	2855	9 ADA01367	Ada01367 Human PRO
35	612	32.1	2855	9 ADA43796	Ada43796 Human CDN
36	612	32.1	2855	9 ADA43564	Ada43564 Human CDN
37	612	32.1	2855	9 ADA01239	Ada01239 Human PRO
38	612	32.1	2855	9 ADA01123	Ada01123 Human CDN
39	612	32.1	2855	9 ADA43680	Ada43680 Human CDN
40	612	32.1	2855	9 ADA06942	Ada06942 Human PRO
41	612	32.1	2855	9 ADA08430	Ada08430 Novel hum
42	612	32.1	2855	9 ADB99723	Adb99723 Human PRO
43	612	32.1	2855	9 ADB87006	Adb87006 Human PRO
44	612	32.1	2855	9 ADB66161	Adb66161 Human CDN
45	612	32.1	2855	10 ADB99839	Adb99839 Human PRO

## ALIGNMENTS

RESULT 1	AAD38695	AAD38695 standard; cDNA; 2369 BP.
ID	AAD38695	
XX		
XX	AAD38695;	
AC		
XX		
DT	23-SEP-2002	(first entry)
XX		
DE	Human LP220	secreted protein encoding cDNA.
XX		
KW	Human; secreted protein; atherosclerosis; Alzheimer's disease; LP220;	
KW	diabetic retinopathy; severe combined immunodeficiency; pancreatitis;	
KW	rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;	
KW	reperfusion injury; arteriosclerosis; wound healing; transgenic animal;	
KW	gene therapy; neoplasia; transgenic; psoriasis; ischaemia; carcinoma;	
XX	chromosome 11q13; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	167..2074
FT		/*tag= a
FT		/product= "Human LP220 secreted protein"
FT	sig_peptide	167..214
FT		/*tag= b
FT	mat_peptide	215..2071
FT		/*tag= c
FT		/product= "Mature human LP220 secreted protein"
XX		
PN	WO200226801-A2.	
XX		
PD	04-APR-2002.	
XX		
PF	14-SEP-2001; 2001WO-US026026.	
XX		
PR	28-SEP-2000; 2000US-0236088P.	
XX		
PA	(ELIL ) LILLY & CO EPL.	
XX		
PI	Su EW, Wang H;	
XX		
DR	WPI; 2002-471259/50.	
DR	P-PSDB; AAE23980.	
XX		





```

Db      847  ||||| CCTGGAGCGGACATCGAGCCATGCTGCTCCCTGACAACCTTCAACTGACATTAACCTT 906
Qy      541  ||||| ATTGACGACTGCCCCCGAGGCGCTTTCGACCACTGGTCACTCTCCGCGCTGACCTC 600
Db      907  ||||| ATTGAGCACTGCCCCCGAGGCGCTTTCGACCACTGGTCACTCTCCGCGCTGACCTC 966
Qy      601  ||||| ACCTCCAAACCGCTGGGCGAGCTGGTCCGGAACCGCTTTTCTCTGTGGGCGTGAATGCA 660
Db      967  ||||| ACCTCCAAACCGCTGGGCGAGCTGGTCCGGAACCGCTTTTCTCTGTGGGCGTGAATGCA 1026
Qy      661  ||||| GAGGCGCTCTCCGCGCGCTGGTGGTGAAGCTTTAGCGGGAACCCCTGCACTGCAACTGT 720
Db      1027  ||||| GAGGCGCTCTCCGCGCGCTGGTGGTGAAGCTTTAGCGGGAACCCCTGCACTGCAACTGT 1086
Qy      721  ||||| GAGCTGTGTGGTGGCGGCGCTGGGCGGAGCGGACGACTGGAAACGTGCGCTCCCG 780
Db      1087  ||||| GAGCTGTGTGGTGGCGGCGCTGGGCGGAGCGGACGACTGGAAACGTGCGCTCCCG 1146
Qy      781  ||||| CCGGCGCTGGCGGCGGCTACTTCTGGGCGAGTCCGAGGGCGAGTTCTCTGTGAGCCG 840
Db      1147  ||||| CCGGCGCTGGCGGCGGCTACTTCTGGGCGAGTCCGAGGGCGAGTTCTCTGTGAGCCG 1206
Qy      841  ||||| CCCCTGATTGCGCGGCAACGACGAGCGCTCTGGGTGTGAGAGGCGCACGCGCTG 900
Db      1207  ||||| CCCCTGATTGCGCGGCAACGACGAGCGCTCTGGGTGTGAGAGGCGCACGCGCTG 1266
Qy      901  ||||| CGGTGCGGCGGCGCTGGGTGACCCCGCGCTACCACTGACTGGGTGGTCTGTGACGACCGG 960
Db      1267  ||||| CGGTGCGGCGGCGCTGGGTGACCCCGCGCTACCACTGACTGGGTGGTCTGTGACGACCGG 1326
Qy      961  ||||| TTGGTTGGCAACTCTCTCCGAGCGCGGGCTTTCCCGCAACGCGGACCTTAGAGATTGGGGTG 1020
Db      1327  ||||| TTGGTTGGCAACTCTCTCCGAGCGCGGGCTTTCCCGCAACGCGGACCTTAGAGATTGGGGTG 1386
Qy      1021  ||||| ACCGCGCTGGGAGACGCTGGGAGCTACACTGTGATGCGCACCAACCTGTGTGAGGCGC 1080
Db      1387  ||||| ACCGCGCTGGGAGACGCTGGGAGCTACACTGTGATGCGCACCAACCTGTGTGAGGCGC 1446
Qy      1081  ||||| ACAGCCCGAGTGAATGCTGGGAGTGTGGCTTGCCTGCTGCTGGGAAACAGAGTGGCGAG 1140
Db      1447  ||||| ACAGCCCGAGTGAATGCTGGGAGTGTGGCTTGCCTGCTGCTGGGAAACAGAGTGGCGAG 1506
Qy      1141  ||||| GGGGGCGCGCGCGCGCGCTGGGACATGCGCGCGCTGCGTGCATGCTGTGCGAGGGTGAG 1200
Db      1507  ||||| GGGGGCGCGCGCGCGCGCTGGGACATGCGCGCGCTGCGTGCATGCTGTGCGAGGGTGAG 1566
Qy      1201  ||||| GGGAGCGTGGAGTCTGAGCCAGCGGTGACGAGTGAACGAGGTGACCGGCACTCAGGCGTG 1260
Db      1567  ||||| GGGAGCGTGGAGTCTGAGCCAGCGGTGACGAGTGAACGAGGTGACCGGCACTCAGGCGTG 1626
Qy      1261  ||||| GTGAGCTGGGGGTCCCGGGCGCGGCGGAGCCGAGTGTGATGTTCCAAATTCAGTACAAAC 1320
Db      1627  ||||| GTGAGCTGGGGGTCCCGGGCGCGGCGGAGCCGAGTGTGATGTTCCAAATTCAGTACAAAC 1686
Qy      1321  ||||| AGCAGCGAAGATGAGACCTCATCTACCGAGATTGTCCAGCGTTCAGGCGCACCACTTCTTG 1380
Db      1687  ||||| AGCAGCGAAGATGAGACCTCATCTACCGAGATTGTCCAGCGTTCAGGCGCACCACTTCTTG 1746
Qy      1381  ||||| CTGAGAGCACTGTGCTCCCGGCGCTGACTATGACTGTGCTGCTGCTGCTGCAACGCGGC 1440
Db      1747  ||||| CTGAGAGCACTGTGCTCCCGGCGCTGACTATGACTGTGCTGCTGCTGCTGCAACGCGGC 1806
Qy      1441  ||||| GCTGGGCGCTTCTGACTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
Db      1807  ||||| GCTGGGCGCTTCTGACTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1866
Qy      1501  ||||| CCGGCGCTGCGCGCGCTGCTGACGCGCGCTGACGAGCGGCGGCGGCGGCGGCGGCGGCGG 1560
Db      1867  ||||| CCGGCGCTGCGCGCGCTGCTGACGCGCGCTGACGAGCGGCGGCGGCGGCGGCGGCGGCGG 1926
Qy      1561  ||||| GCGCGTGGGCGGCGTGTGCTGTGAGCTGACTTACTGTGCTTCACTGTGCGCTTGTGCTGCG 1620

```

```

Db      1927  ||||| GCGGTGGGCGGCGTGTGCTGAGTGGCTGCTACTGCTTCACTGTGAGCTGTGCTGCG 1986
Qy      1621  ||||| GCGCGGGGGGCGGGAATGAGCGCGCTCCCGCTTAAGCTCAGCGCACTCCAGTCCGAGCC 1680
Db      1987  ||||| GCGCGGGGGGCGGGAATGAGCGCGCTCCCGCTTAAGCTCAGCGCACTCCAGTCCGAGCC 2046
Qy      1681  ||||| AATGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1740
Db      2047  ||||| AATGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2106
Qy      1741  ||||| CAGCGAGCTGCTCTCTGGAAGTGGAGATGCGGGTGTCTACGTTATGCGAGCGCGCTG 1800
Db      2107  ||||| CAGCGAGCTGCTCTCTGGAAGTGGAGATGCGGGTGTCTACGTTATGCGAGCGCGCTG 2166
Qy      1801  ||||| GGAAGAGCTTTGGGCGCGGCGGAGCCACTCTGTGTCAATGGGGGCGTGTCTCGGGGCGAGGTC 1860
Db      2167  ||||| GGAAGAGCTTTGGGCGCGGCGGAGCCACTCTGTGTCAATGGGGGCGTGTCTCGGGGCGAGGTC 2226
Qy      1861  ||||| CCGGGGGTGAAGAGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1908
Db      2227  ||||| CCGGGGGTGAAGAGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2274

```

## RESULT 3

AD807450  
ID ADE07450 standard; DNA; 2364 BP.

AC ADE07450;

XX 29-JAN-2004 (first entry)

DE Novel coding sequence (useful for identifying genetic disorders) #516.

XX novel gene; novel protein; tissue marker; molecular weight marker;

KM chromosome marker; genetic disorder; gene; ds.

OS Unidentified.

PN MO2003054152-A2.

XX 03-JUL-2003.

XX 10-DEC-2002; 2002MO-US039555.

XX 10-DEC-2001; 2001US-0339739P.

PR 11-DEC-2001; 2001US-0339453P.

PR 14-MAR-2002; 2002US-0365091P.

PR 12-MAR-2002; 2002US-0365384P.

PR 12-APR-2002; 2002US-0372615P.

PR 22-APR-2002; 2002US-00128558.

PR 24-APR-2002; 2002US-0376045P.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J,

PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z,

PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ,

XX MPI, 2003-569235/53.

XX P-PSDB; ADE08361.

XX New polynucleotides, useful for expressing recombinant proteins for

PT analysis, characterization or therapeutic use, or as markers for tissues

XX in which the corresponding protein is preferentially expressed.

XX Claim 1; SEQ ID NO 516; 1177bp; English.

XX The invention comprises the amino acid and coding sequences of novel

CC proteins. The DNA and protein sequences of the invention are useful as:

CC markers for tissues in which the corresponding protein is preferentially

CC expressed; as molecular weight markers on gels; as chromosome markers or

CC tags; to identify chromosomes or to map related gene positions; and to





```

XX Human NOV44a coding sequence.
DE
XX
XX Human; anti-HIV; cytostatic; antidiabetic; antislutmatic; cachexia; AIDS;
XX antiinflammatory; cardiac; haemostatic; neuroprotective; anorectic;
XX neurotropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
XX antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
XX metabolic disorder; diabetes; obesity; infectious disease; anorexia;
XX neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
XX immune disorder; haematopoietic disorder; cardiovascular disorder;
XX bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
XX metabolic syndrome X; wasting disorder; cell differentiation;
XX Single nucleotide polymorphism; SNP; cell proliferation; haematopoiesis;
XX wound healing; angiogenesis; gene; ds.
XX
OS Homo sapiens.
XX
XX Key
XX Location/Qualifiers
XX variation
XX /tag= a
XX replace(604,G)
XX
XX variation
XX /standard_name= "Single nucleotide polymorphism"
XX replace(684,G)
XX /tag= b
XX /standard_name= "Single nucleotide polymorphism"
XX replace(1123,A)
XX
XX variation
XX /tag= c
XX /standard_name= "Single nucleotide polymorphism"
XX replace(1489,C)
XX /tag= d
XX /standard_name= "Single nucleotide polymorphism"
XX
XX
XX WO200272771-A2.
XX
XX 19-SEP-2002.
XX
XX 08-MAR-2002; 2002WO-US007288.
XX
XX
XX 08-MAR-2001; 2001US-0274101P.
XX 08-MAR-2001; 2001US-0274194P.
XX 08-MAR-2001; 2001US-0274281P.
XX 08-MAR-2001; 2001US-0274322P.
XX 09-MAR-2001; 2001US-0274849P.
XX 12-MAR-2001; 2001US-0275235P.
XX 13-MAR-2001; 2001US-0275578P.
XX 13-MAR-2001; 2001US-0275579P.
XX 13-MAR-2001; 2001US-0275601P.
XX 14-MAR-2001; 2001US-0276000P.
XX 16-MAR-2001; 2001US-0276776P.
XX 19-MAR-2001; 2001US-0276994P.
XX 20-MAR-2001; 2001US-0277239P.
XX 20-MAR-2001; 2001US-0277321P.
XX 20-MAR-2001; 2001US-0277327P.
XX 20-MAR-2001; 2001US-0277387P.
XX 21-MAR-2001; 2001US-0277791P.
XX 22-MAR-2001; 2001US-0277833P.
XX 23-MAR-2001; 2001US-0278152P.
XX 26-MAR-2001; 2001US-0278894P.
XX 27-MAR-2001; 2001US-0278999P.
XX 27-MAR-2001; 2001US-0279036P.
XX 28-MAR-2001; 2001US-0279344P.
XX 30-MAR-2001; 2001US-0279955P.
XX 30-MAR-2001; 2001US-0280233P.
XX 02-APR-2001; 2001US-0280802P.
XX 02-APR-2001; 2001US-0280822P.
XX 02-APR-2001; 2001US-0280900P.
XX 04-APR-2001; 2001US-0281194P.
XX 13-APR-2001; 2001US-0283675P.
XX 30-APR-2001; 2001US-0283742P.
XX 03-MAY-2001; 2001US-0288066P.
XX 03-MAY-2001; 2001US-0288342P.
XX 03-MAY-2001; 2001US-0288528P.
XX 15-MAY-2001; 2001US-0291190P.
XX 16-MAY-2001; 2001US-0291099P.

```

```

PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332712P.
PR 14-NOV-2001; 2001US-0332722P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333722P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 08-MAR-2002; 2002US-00093463.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ,
XX Boldog FL, Li L, Zehusen BD, Tehernev VT, Gangoli EA, Vernet CM;
XX Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spiderna SK,
XX Voss EZ, Malyanar UM, Anderson DW, Parturajan M, Miller CE,
XX Taupier RJ, Padigaru W, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
XX Zhong M;
XX
XX WPI; 2002-732824/79.
XX P-FSDB; ABP70142.
XX
XX DR
XX
XX PT New NOVX polypeptides and polynucleotides, useful for preventing,
XX diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
XX Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
XX disorders, and asthma.
XX
XX PS Claim 16; Page 261-262; 619pp; English.
XX
XX CC The present invention relates to new isolated proteins (NOVX) and their
XX coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
XX any number from 1 to 48. The NOVX proteins and coding sequences are
XX useful in the manufacture of a medicament for treating a syndrome
XX associated with a human disease, preferably a NOVX-associated disorder.
XX The NOVX coding sequences and proteins are useful for treating, diabetes,
XX preventing or diagnosing diseases such as metabolic disorders, diabetes,
XX obesity, infectious disease, anorexia, cancer-associated cachexia,
XX cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
XX disease, immune disorders, haematopoietic disorders, cardiovascular
XX disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
XX disturbances associated with obesity, metabolic syndrome X or wasting
XX disorders associated with chronic diseases or various cancers. The NOVX
XX coding sequences and proteins may also be used as targets for the
XX identification of small molecules that modulate or inhibit e.g.
XX neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
XX wound healing and angiogenesis, in gene therapy, in generation of
XX antibodies that bind immunospecifically to NOVX substances for use in
XX therapeutic or diagnostic methods
XX
XX Sequence 3201 BP; 534 A; 1174 C; 986 G; 507 T; 0 U; 0 Other;

```

```

Query Match 99.4%; Score 1897; DB 6; Length 3201;
Best Local Similarity 99.98; Pred. No. 0;
Matches 1908; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```



OS Homo sapiens.  
 XX MO2003025148-A2.  
 PN 27-MAR-2003.  
 XX 19-SEP-2002; 2002MO-US029964.  
 PF 19-SEP-2001; 2001US-0323739P.  
 PR 13-SEP-2002; 2002US-00323739.  
 XX (HXSE-) HXSEQ INC.  
 PA Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;  
 PI Ghosh M, Xue AJ, Weinman T, Weng G, Zhou P, Drmanac RT, Wang D;  
 PI Haley-Vicente D;  
 XX WPI; 2003-354603/33.  
 DR P-PSDB; ADI21104.  
 XX New polynucleotides and secreted proteins, useful for treating myeloid or  
 PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve  
 PT tissue growth or regeneration, in wound healing, and in tissue repair and  
 PT replacement.  
 XX Claim 1; SEQ ID NO 79; 156bp; English.  
 PS The invention relates to an isolated polynucleotide encoding a  
 XX polypeptide with biological activity. The polynucleotides and  
 CC polypeptides are useful in diagnostics, forensics, gene mapping,  
 CC identification of mutations responsible for genetic disorders and other  
 CC traits, to assess biodiversity, as nutritional sources or supplements.  
 CC The polynucleotides may also be used as molecular weight markers,  
 CC chromosome markers or map related gene positions, or as an antigen to  
 CC raise anti-DNA antibodies or elicit immune response. The polypeptides are  
 CC useful for raising antibodies, as markers for tissues in which the  
 CC corresponding polypeptide is expressed, for re-engineering damaged or  
 CC diseased tissues, for treating myeloid or lymphoid cell disorders, in  
 CC bone cartilage, tendon, ligament and/or nerve tissue growth or  
 CC regeneration, in wound healing, in tissue repair and replacement, in  
 CC healing of burns, incisions and ulcers, and in treating cancer. The  
 CC present sequence represents cDNA encoding a novel human protein.  
 CC  
 XX Sequence 2885 BP; 443 A; 989 C; 930 G; 523 T; 0 U; 0 Other;  
 SQ  
 Query Match 93.8%; Score 1789.6; DB 10; Length 2885;  
 Best Local Similarity 97.1%; Pred. No. 6e-306;  
 Matches 1853; Conservative 0; Mismatches 4; Indels 51; Gaps 1;

QY 361 AATCTGACGACCTCATCTCAGCGGCAACAGCTGGGCGGCACTCGCGGAGCCTTC 420  
 |||||  
 Db 1097 AATCTGACGACCTCATCTCAGCGGCAACAGCTGGGCGGCACTCGCGGAGCCTTC 1156  
 QY 421 GACGACTTCTTAGAGAGCTGAGAGACTTGAGACTGTCTTACAACAACCTCGGCAAGTG 480  
 |||||  
 Db 1157 GACGACTTCTTAGAGAGCTGAGAGACTTGAGACTGTCTTACAACAACCTCGGCAAGTG 1216  
 QY 481 CCTTGGGCGGCAATGGCGGCAATGCTGCTGCTGCAACCTTCAACCTGAGCACTTAACCTT 540  
 |||||  
 Db 1217 CCTTGGGCGGCAATGGCGGCAATGCTGCTGCTGCAACCTTCAACCTGAGCACTTAACCTT 1276  
 QY 541 ATTGACGACCTGCGCCGACGAGCGCTTTCGACAGCTCGGTGAGCTTCCGCTGAGCTTC 600  
 |||||  
 Db 1277 ATTGACGACCTGCGCCGACGAGCGCTTTCGACAGCTCGGTGAGCTTCCGCTGAGCTTC 1336  
 QY 601 ACCTTCAACCGCTGACGCTGACGCTGCTGCGGACCGCTTTCTCTGCGGCGGTATGCA 660  
 |||||  
 Db 1337 ACCTTCAACCGCTGACGCTGACGCTGCTGCGGACCGCTTTCTCTGCGGCGGTATGCA 1386  
 QY 661 GAGGCGCTTCGCGCGCGCGCGCTGAGCTTGAAGGGGAACCCCTGCACTGCACTGT 720  
 |||||  
 Db 1387 -----AACCTTGAAGCTGCACTGT 1405  
 QY 721 GAGCTGCTGAGCTGCGGCGGCTGAGCGCGGACGAGCACTGGAACGTGCGCTTCCCG 780  
 |||||  
 Db 1406 GAGCTGCTGAGCTGCGGCGGCTGAGCGCGGACGAGCACTGGAACGTGCGCTTCCCG 1465  
 QY 781 CCGGCGCTGCGCGCGCTTACTTCTGAGGAGTGCCGAGGCGAGTTCTCTGTGAGCCG 840  
 |||||  
 Db 1466 CCGGCGCTGCGCGCGCTTACTTCTGAGGAGTGCCGAGGCGAGTTCTCTGTGAGCCG 1525  
 QY 841 CCGCTCATTCGCGCGCAACAGAGCGCTTGGAGTGTTGAAGCGGCGGCGCAAGCTG 900  
 |||||  
 Db 1526 CCGCTCATTCGCGCGCAACAGAGCGCTTGGAGTGTTGAAGCGGCGGCGCAAGCTG 1585  
 QY 901 CCGTGCCTGCGGCGCTTGGAGTGACCCGCGCTTACCATGACATGAGTGGTCTTGAAGCGG 960  
 |||||  
 Db 1586 CCGTGCCTGCGGCGCTTGGAGTGACCCGCGCTTACCATGACATGAGTGGTCTTGAAGCGG 1645  
 QY 961 TTGGTTGGCAACTCTTCGCGAGCGCGGCTTTCGCAACGCGGACCTTGAAGATTGGAGTG 1020  
 |||||  
 Db 1646 TTGGTTGGCAACTCTTCGCGAGCGCGGCTTTCGCAACGCGGACCTTGAAGATTGGAGTG 1705  
 QY 1021 ACCGCGCTTGGAGAGCACTGAGGAGCTTACACTGATGCGCAACCTGCTGAGAGGCC 1080  
 |||||  
 Db 1706 ACCGCGCTTGGAGAGCACTGAGGAGCTTACACTGATGCGCAACCTGCTGAGAGGCC 1765  
 QY 1081 ACAGCCGAGTAGAAGCTGCGAGTGTGACCTTGGCCCATGATGATGGAGAACAGACAGTCCGAG 1140  
 |||||  
 Db 1766 ACAGCCGAGTAGAAGCTGCGAGTGTGACCTTGGCCCATGATGATGGAGAACAGACAGTCCGAG 1825  
 QY 1141 GGGGCGCGCGCGCGCGCTTGGAGCATGCGCGCTTCCGCTGAGCTGTGCGGAGAGGTGAG 1200  
 |||||  
 Db 1826 GGGGCGCGCGCGCGCGCTTGGAGCATGCGCGCTTCCGCTGAGCTGTGCGGAGAGGTGAG 1885  
 QY 1201 GGGAGCTGGAAGTCTGAGAGCAAGCGGTGAGAGTGAAGAGTGAACGCGCACTTCAAGGAGCTG 1260  
 |||||  
 Db 1886 GGGAGCTGGAAGTCTGAGAGCAAGCGGTGAGAGTGAAGAGTGAACGCGCACTTCAAGGAGCTG 1945  
 QY 1261 GTGAGCTGAGGAGTCCGCGGCGGCGGAGCGGACCGAGCTGATGATGTTCCAAATCCAGTAAAC 1320  
 |||||  
 Db 1946 GTGAGCTGAGGAGTCCGCGGCGGCGGAGCGGACCGAGCTGATGATGTTCCAAATCCAGTAAAC 2005  
 QY 1321 AGCAGGGAAGATGAGAACCTTCACTTACCGGATTTGTCCAGGCTTCCAGGCACTTCCCTG 1380  
 |||||  
 Db 2006 AGCAGGGAAGATGAGAACCTTCACTTACCGGATTTGTCCAGGCTTCCAGGCACTTCCCTG 2065  
 QY 1381 CTGAAGCACTCTTCCCGCGCGCTGACTATGACCTTCTGCTGTGAGCTTGTACCGGCGC 1440  
 |||||  
 Db 2066 CTGAAGCACTCTTCCCGCGCGCTGACTATGACCTTCTGCTGTGAGCTTGTACCGGCGC 2125  
 QY 1441 GCTTGGGCGCTTGAAGCTCAAGGCGCACAGGCTGCTGAGCTTGTCCCACTTCTCCAGGCTG 1500  
 |||||









QY 901 CGGTGCGGGGCGCTGGGTGACCCCGGCGCTACCATGCACTGGGTGGTCTCTGACGACCGG 960  
DB 901 CGGTGCGGGGCGCTGGGTGACCCCGGCGCTACCATGCACTGGGTGGTCTCTGACGACCGG 960  
QY 961 TTGTTGGGCACTCTCCCGAGCGCGGGCTTTCCCGCAAGGGAAGCTTAGAGATTGGGGTGG 1020  
DB 961 TTGTTGGGCACTCTCCCGAGCGCGGGCTTTCCCGCAAGGGAAGCTTAGAGATTGGGGTGG 1020  
QY 1021 ACCGCGCTGGGGGCGCTGGGGGCTACACCTGATGCGACCAACCTTGTGGTGAAGCC 1080  
DB 1021 ACCGCGCTGGGGGCGCTGGGGGCTACACCTGATGCGACCAACCTTGTGGTGAAGCC 1080  
QY 1081 ACGAGCGGAGTGAATGCGGGGTGGTGGCTTGGCCCAATGGTGGGAAAGCGAGTGGCCGAG 1140  
DB 1081 ACGAGCGGAGTGAATGCGGGGTGGTGGCTTGGCCCAATGGTGGGAAAGCGAGTGGCCGAG 1140  
QY 1141 GGGGGCGGGCGCGGCGCTCGAGCATGCGCGCTCGCTCGCATGCTGCGGAGGGTGGAG 1200  
DB 1141 GGGGGCGGGCGCGGCGCTCGAGCATGCGCGCTCGCTCGCATGCTGCGGAGGGTGGAG 1200  
QY 1201 GGGAGCGCTGAGTCTGAGCCAGCGGTGCAAGTGAAGGAGTGAAGCGGCACTGAGGGCTGG 1260  
DB 1201 GGGAGCGCTGAGTCTGAGCCAGCGGTGCAAGTGAAGGAGTGAAGCGGCACTGAGGGCTGG 1260  
QY 1261 GTGAGCTGGGGGTCCCGGGGGCGGCGAGCGGCGGAGTGTGATGTTCCAAATCCAGTGAAC 1320  
DB 1261 GTGAGCTGGGGGTCCCGGGGGCGGCGAGCGGCGGAGTGTGATGTTCCAAATCCAGTGAAC 1320  
QY 1321 AGCAGCGAAGATGAGACCCCTCATCTACCGGATTTGCTCCAGCGCTCGAGCAACACTTCTCG 1380  
DB 1321 AGCAGCGAAGATGAGACCCCTCATCTACCGGATTTGCTCCAGCGCTCGAGCAACACTTCTCG 1380  
QY 1381 CTGAGAGCACTCTGCTCCCGGCGCTGACATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
DB 1381 CTGAGAGCACTCTGCTCCCGGCGCTGACATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
QY 1441 GCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
DB 1441 GCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
QY 1501 CCGGCGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
DB 1501 CCGGCGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
QY 1561 GCGGCGGGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1618  
DB 1561 GCGGCGGGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1618  
QY 1619 GGGGCGGGGGGGCGGGAATGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1651  
DB 1621 GGGGCGGGGGGGCGGGAATGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1651

RESULT 8  
ABV99421  
ID ABV99421 standard; DNA; 2451 BP.  
AC ABV99421;  
XX 27-JAN-2003 (first entry)  
XX  
DE Human NOV44b coding sequence.

KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.  
XX Homo sapiens.  
OS  
XX  
PN WO200272771-A2.  
PD 19-SEP-2002.  
XX  
PF 08-MAR-2002; 2002WO-US007288.  
XX  
PR 08-MAR-2001; 2001US-0274101P.  
PR 08-MAR-2001; 2001US-0274194P.  
PR 08-MAR-2001; 2001US-0274281P.  
PR 08-MAR-2001; 2001US-0274322P.  
PR 09-MAR-2001; 2001US-0274489P.  
PR 12-MAR-2001; 2001US-0275235P.  
PR 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
PR 13-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
PR 16-MAR-2001; 2001US-0276766P.  
PR 19-MAR-2001; 2001US-0276994P.  
PR 20-MAR-2001; 2001US-0277239P.  
PR 20-MAR-2001; 2001US-0277321P.  
PR 20-MAR-2001; 2001US-0277327P.  
PR 20-MAR-2001; 2001US-0277338P.  
PR 21-MAR-2001; 2001US-0277919P.  
PR 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
PR 26-MAR-2001; 2001US-0278894P.  
PR 27-MAR-2001; 2001US-0278999P.  
PR 27-MAR-2001; 2001US-0279036P.  
PR 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0279995P.  
PR 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280802P.  
PR 02-APR-2001; 2001US-0280822P.  
PR 02-APR-2001; 2001US-0281194P.  
PR 04-APR-2001; 2001US-0281099P.  
PR 13-APR-2001; 2001US-0283675P.  
PR 30-APR-2001; 2001US-0287424P.  
PR 02-MAY-2001; 2001US-0288066P.  
PR 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288528P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
PR 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294889P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 18-JUN-2001; 2001US-0294899P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.  
PR 12-SEP-2001; 2001US-0318770P.  
PR 27-SEP-2001; 2001US-0325430P.  
PR 27-SEP-2001; 2001US-0325681P.  
PR 18-OCT-2001; 2001US-0330380P.  
PR 31-OCT-2001; 2001US-0335301P.  
PR 14-NOV-2001; 2001US-0332172P.  
PR 14-NOV-2001; 2001US-0332271P.  
PR 14-NOV-2001; 2001US-0332272P.  
PR 14-NOV-2001; 2001US-0333184P.  
PR 14-NOV-2001; 2001US-0333272P.  
PR 21-NOV-2001; 2001US-0332094P.  
PR 03-DEC-2001; 2001US-0337426P.  
PR 03-DEC-2001; 2001US-0338092P.  
PR 04-DEC-2001; 2001US-0337185P.  
PR 03-JAN-2002; 2002US-0345705P.  
PR 08-MAR-2002; 2002US-00093463.









Db 121 TTYGTCNCNNAAYGTNGAYMGMNACNGTNGARYTNMGTNTGNCNAAAYATTYATH 180  
 QY 181 CAGGCCCTGGGGCCCCCTGACTTCCCAACATGACGGGACTGTGACCTGACACTGTCT 240  
 Db 181 CAGGCTTNGGNCNCNGAYTTTGMNAAVATGACNGAYTNGTNGAYTTNACNTNMN 240  
 QY 241 CGCAATGCCATCAACCGGCAATGGGGCCCGCTTTGGGGAACCTCGAAGACCTCGGTCC 300  
 Db 241 MGNAAVGCNATHACNNGNATHGNGNCNNGCNTTYTGNGAYTTNGARFMSNTNMN 300  
 QY 301 CTCACCTTGACGGCAACAGGCTGTGAGCTGGGCAACGGGACCTCGGGGCCCCGTC 360  
 Db 301 YTNCAAYTNGAYGNAAYMGNVNTNGNARYTNGNACNGMWSNTNMNGNCGN 360  
 QY 361 AATTCGACACCTATCTCTCAGCGCAACAGCTGGGCGCATCGCGCGGAGCCTTC 420  
 Db 361 AAYYTNCAACAYATYNAHYTNMWSNGNAAVCARYTNGMGNATHGNCNCGNCGNTTY 420  
 QY 421 GACGACTTCTAGAGAGCTGGAGACCTGGACCTGCTCTACAAACAACCTCCGCAAGTG 480  
 Db 421 GAYGATTTTNGARMSNTNGARAYTNGAYTNMWSNTYAAAYATNMGCARGTN 480  
 QY 481 CCTGGGCGGACATGGGCGCATGCTGCTGCAACCTGACCTGACCATACCTT 540  
 Db 481 CCGTGGGNGNATHGNGNATGCGNGNTNCAVACNTMAAYTTNGAYTAAAYTN 540  
 QY 541 ATTGACGACTGCCCCCAAGGCGCTTGCCCAAGCTCGGTGACCTCCGCGCTGACCTC 600  
 Db 541 ATHGAYGNTNCCNCGNGGNCNTTYGNCARNTNGNCARNTMWSMGNVNTNGAYTN 600  
 QY 601 ACCTCAACCGGCTGGGACGCTGGCTCGGACCCGCTTTCTCTGCTGGGCGTGATCA 660  
 Db 601 ACNMSNAAVMTYNTNGNACNNTNGCNCNAYCCNNTNTTWSMNGMNGMNGAYG 660  
 QY 661 GAGGCTCTCCGCCCCCTGCTGCTGAGCTTTAGCGGAAACCCCTGCACTGAACTGT 720  
 Db 661 GARGGNMNCNCGNCNCTYNTGNTTMSNTTYSNGNAAVCCNNTNCAITGTAAYTY 720  
 QY 721 GAGCTGCTGTGCTGGCGGCGCTGGCGGCGGACGACGACTGGAACGCTGCGCTCCCG 780  
 Db 721 GARTYNTNTGGYTNMGMNNTNGCNMNGNCNAGAYYTNGARACNTGCGMWSN 780  
 QY 781 CCGGCGCTGGCGGCGCTGCTGCTGCACTGCTGCGGCGGCTGCTGCTGAGCGG 840  
 Db 781 CCGGNTTNGCNGMNTATYTTGCGCNGTNCNARGNCGRTYWSNTYGARCN 840  
 QY 841 CCCCCTCATTTGCCGCAACGACGAGCGCTGCTGCTGGAAGGCGGCGGACGCTG 900  
 Db 841 CCNNTNATHGCMNGNCAVACNARMGNTTGGTNTNGARSGNCAARMGNCNACNTN 900  
 QY 901 CGGTGGCGGCGCTGGGTGACCCCGGCTTACCATGCACTGGGTGGTCTGACGACG 960  
 Db 901 MGNITGNGCNGTNTNGNAGYCCNGCNCNACNATGATYGGGTNGCNGAYGAMGN 960  
 QY 961 TTGGTTGGCAACTCTCTCCGAGCCCGGCTTCCCGCAACGGGACCTTGAAGATTGGG 1020  
 Db 961 YTNGTNGNAAVMSNMSMNGCMNGCNVTCNNAAYGNCNNTNNGARHTHGAGTN 1020  
 QY 1021 ACCGGCGCTGGGAGCGCTGCGGCGCTACACTGATCGCCCAACCTGCTGAGGGCC 1080  
 Db 1021 ACNGGNGCNGNGAYGNGNGNTYACNTGYATHGNCNAAVCCNGCNGNAGAGCN 1080  
 QY 1081 ACAGCCCGAGTAACTGCGGAGTGTGGGCTTGCCCCATGTTGGGAACAGCAGTGGCG 1140  
 Db 1081 ACNGCMNGNTNGARYTNMGNNTNNGCNVTCNCAVAGNGNAAVMSMNGCNGAR 1140  
 QY 1141 GGGGGCGGCGGCGCTCGACATCGCGGCTCCGCTGCACTGCTGCCAGGCTGAG 1200  
 Db 1141 GGGGMMNGCNGNCGNMSNGAYATHGNCNGMWSMNGNACNCGNCGNARGAGAR 1200  
 QY 1201 GGGAGCTGAGTCTGAGCCAGCGCTGACAGTGTGACGAGGTGACCCCACTGAGGCG 1260  
 Db 1201 GGNACVYTNGARFMSNGARCCNGCNGTNGARNTNACNGARNTNACNCGMWSNGAYTN 1260

QY 1261 GTGAGCTGGGCTCCCGGGGCGGCGGACCCAGCTGTGATGTTCCAAATCCAGTACAC 1320  
 Db 1261 GTMWSNTGGGNCNCGMNGNCNCGNAYCCNNTNGATGTTGACARATHCARAYVAY 1320  
 QY 1321 ACAGCGAAGATGAGACCTCTACCGGATTTGCCAGCTCCAGCCACCACTTCCTG 1380  
 Db 1321 MSNMNGARAGAYGABACNTYNTATHMGNATHGNCNGMWSMNCAYCAVTTYTN 1380  
 QY 1381 CTGAAGCACTGTCTCCCGGCGCTGACTATGACCTCTGCTGCTGCTGCTGCTGCTG 1440  
 Db 1381 YTNARCAVYTNGTNCNCGNCGNCAVAYGAYTNTGYTNNTGNTNMNCGN 1440  
 QY 1441 GCTGGGCTCTGACTCTACCGGCAACAGGCTGCTGGCTGTGCCATTTCTTCAAGCTG 1500  
 Db 1441 GNGNCCNWSNGAYYTNACNCGNACNNGNNTYNTNGNTGVCNCAVTTYWSNACNTN 1500  
 QY 1501 CCGGCTGCGCCCTGTGCTGCGGCGCTGAGGCGGCTGCTGAGCCCTGACGCTG 1560  
 Db 1501 CCGMWSNCCNNTNTGCAVGCNTNCAAGCNCAYGNTNNGGNGNACNTNACNTN 1560  
 QY 1561 GCGGTGGGGGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1618  
 Db 1561 GCGTNGGNGGNTYNTGTCNCGNCAVYTNACNCGNNTNCAVYTGCGMCMNTGYTGT 1620  
 QY 1619 GGGGCGGCGGCGGCGGAAATGCGCGCTCCCG 1650  
 Db 1621 GNGCNGGNGGNCNCGARATGCGNCGMWSNCC 1652

RESULT 12  
 AA166986  
 ID AA166986 standard; cDNA; 1887 BP.  
 AC AA166986;

30-JAN-2002 (first entry)

Human LRR family member, 33395 coding sequence.

LRR; leucine rich repeat; 33395; cytoskeletal; anti-HIV; antidiabetic;  
 antiarthritic; neuroprotective; dermatological; immunosuppressive; ss;  
 antiinflammatory; antislathmatic; antitumor; antianaemic; hepatotropic.  
 OS Homo sapiens.

FH Key Location/Qualifiers  
 FT 1..1887  
 FT /\*tag= a  
 FT /product= "33395"  
 PN W0200172827-A2.

04-OCT-2001.

23-MAR-2001; 2001WO-US009470.

24-MAR-2000; 2000US-0191863P.

(MILL-) MILLENNIUM PHARM INC.

Glucksmann MA;

WPI; 2001-626254/72.

P-PDB; AAG65805.

New polypeptide for preventing or treating disorders associated with  
 cellular adhesion, proliferation or differentiation, comprises  
 polypeptide 33395, a member of the leucine rich repeat protein family.  
 Claim 1, Fig 1; 133pp; English.

The invention provides an isolated nucleic acid encoding a polypeptide of





1278 TACCGACCGTGGCGTCCAGGTGACTGAGCACCAGGAGCCACAGTGTCTTCTCCAGTCCG 1337

Db 1278 TACCGACCGTGGCGTCCAGGTGACTGAGCACGGGCCACAGCTGCTTGTGCCAGTGGCC 1337

QY 1272 TCCGGGCGGCGCAGCCAGCCAGCTGTGATGTTCCAAATCCAGTACAAAGCAGCAAGA 1331  
 DB 1338 GGAATCAGCGGCTATCCCGGGATCCGCAATGTACCAAGATCCAGTACAAAGCAGCTGTG 1397  
 QY 1332 TGAGACCTCTATTTACCGGATTTGCCAGCTCCAGCCACCACTTCTGTGTGAAGACCT 1391  
 DB 1398 TGACATCTCTGTCTACAGATGATCCCGGGAGAGCGCTGCTCTGTGTGAAGACCT 1457  
 QY 1392 CTTCCCGGCGGCTGACATGATGATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1451  
 DB 1458 GGGCTAGCGCGGACCTTACATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1514  
 QY 1452 TGACCTCAGCGGCGCAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1511  
 DB 1515 GGGCTCAGCGGCGCAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1574  
 QY 1512 CTTGTGCGACGCGCTGT 1571  
 DB 1575 GCGATGTGCGGCGCGCGACGCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1634  
 QY 1572 TGTGCTGT 1618  
 DB 1635 CGTATCTGT 1694  
 QY 1619 --GGGGCGGCGGCGGCGGAAATGGCGCGCTCTCTGTGTGTGTGTGTGTGTGTGT 1676  
 DB 1695 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1754  
 QY 1677 GACCAATGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1718  
 DB 1755 GACCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1796

RESULT 14  
 ID AAH25568 standard; DNA, 1887 BP.  
 AC AAH25568;  
 XX  
 DT 05-SEP-2001 (first entry)

Nucleotide sequence of an interferon omega-1 like protein NOV2.  
 DE Interferon omega-1 like protein; NOV2; membrane bound protein;  
 XX secreted protein; spermatogenesis; male infertility; neoplasia;  
 XX blood circulation; immunological disorder; autoimmune disease;  
 XX inflammatory disease; cardiovascular disease; metabolic disease; cancer;  
 XX viral infection; acute lymphoblastic leukemia; glioma;  
 XX neurological disease; neurodegenerative disorder; Alzheimer's disease;  
 XX Parkinson's disorder; hematopoietic disorder; neurodegenerative disorder;  
 XX immune disorder; hematopoietic disorder; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT 1.1887  
 FT CDS /product= "interferon omega-1 like protein NOV2"  
 FT  
 XX MO200142471-A2.  
 XX  
 XX 14-JUN-2001.  
 PD  
 PF 08-DEC-2000; 2000MO-US033463.  
 XX  
 PR 09-DEC-1999; 99US-0169887P.  
 PR 10-DEC-1999; 99US-0170230P.  
 PR 07-DEC-2000; 2000US-00732436.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 PI Prayaga SK, Shimkets RA;

XX  
 DR WPI: 2001-381691/40.  
 DR P-PSDB; AAB84469.  
 XX  
 PT Novel polypeptides designated as NOV polypeptides, useful in detection,  
 PT prevention and treatment of a broad range of pathological states.  
 XX  
 PS Claim 8; Page 15-16; 92pb; English.  
 XX  
 CC The present sequence encodes an interferon omega-1 like protein,  
 CC designated NOV2. The protein is a membrane bound and secreted protein.  
 CC NOV polypeptides and polynucleotides are useful for treating of a  
 CC preventing a NOV-associated disorder, and in the manufacture of a  
 CC medicament for treating a syndrome associated with a human disease such  
 CC as NOV-associated disorder. They are also useful for treating disorders  
 CC or syndromes including those involved in development, differentiation and  
 CC activation of thymic immune cells, pathologies related to spermatogenesis  
 CC and male infertility, diagnosis of blood circulation and/or pathologies, in disease  
 CC and pathologies of cells in blood circulation and/or pathologies, in disease  
 CC platelets, various immunological disorders and/or pathologies, in disease  
 CC and inflammatory diseases, cardiovascular diseases, metabolic diseases,  
 CC cancer growth and metastasis, viral infections, cancer therapy, acute  
 CC lymphoblastic leukemia, in gliomas, neurological diseases,  
 CC neurodegenerative disorders, Alzheimer's disease, Parkinson's disorder,  
 CC and hematopoietic disorders. NOV polypeptides and polynucleotides are  
 CC also useful for treating neurodegenerative disorders, immune disorders  
 CC and hematopoietic disorders  
 XX  
 SO Sequence 1887 BP; 259 A; 736 C; 577 G; 315 T; 0 U; 0 Other;

Query Match 32.7%; Score 624.8; DB 4; Length 1887;  
 Best Local Similarity 62.4%; Pred. No. 6.4e-101;  
 Matches 1074; Conservative 0; Mismatches 597; Indels 51; Gaps 4;

QY 45 CGCTGCGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 104  
 DB 78 CCGATGCTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 137  
 QY 105 CCACCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 164  
 DB 138 AGGGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 197  
 QY 165 GGCTGCAACTTCAATCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 224  
 DB 198 GGCAGCAACTTCAATCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 257  
 QY 225 GGACCTGCAACTTCAATCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 284  
 DB 258 GCACTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 317  
 QY 285 CGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 344  
 DB 318 GGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 377  
 QY 345 CCTCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404  
 DB 378 GCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437  
 QY 405 CGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 464  
 DB 438 GGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 497  
 QY 465 CAACTTCCGAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 524  
 DB 498 CAACCTTCCGAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 557  
 QY 525 CTTGAGACATTAATTTAGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 584  
 DB 558 CTTGAGACATTAATTTAGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 617  
 QY 585 CTGCGGCTGAGACTCACTTCAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 644  
 DB 618 GGGCGGCTGAGACTCACTTCAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 677

QY 645 TCGTGGG-----GTATGACAGAGGCTCTCCGCGCCCTGTGTGACTT 692  
DB 678 CCGCTGCCCCCTGCTGCGAGGCCCCGAGGCTGCGCCCTGCTGCTGCTGCTGCTGCT 737  
QY 693 TAGCGGGAACCCCTGCACTGCACTGTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 752  
DB 738 TGGCGGGAGCCCTGCACTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 797  
QY 753 GAGCAGCTGTGAAGAGT 812  
DB 798 GAGCAGCTGT 857  
QY 813 GCGCGAGGCGGAGT 872  
DB 858 GAGCAGAGGAGGAGT 917  
QY 873 GGTGTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 932  
DB 918 TGTGCCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 977  
QY 933 CATGACTGT 992  
DB 978 TGT 1037  
QY 993 CCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1052  
DB 1038 CCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1097  
QY 1053 CATGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1103  
DB 1098 CATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1157  
QY 1104 -----GCTGCTTGT 1151  
DB 1158 ACGACCTGT 1217  
QY 1152 CCGGCGCTGT 1211  
DB 1218 TGTCTTACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1277  
QY 1212 GTGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1271  
DB 1278 TACGACCTGT 1337  
QY 1272 TCCCGGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1331  
DB 1338 GATCAGCGGCTATCCCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1397  
QY 1332 TAGACCTGT 1391  
DB 1398 TGTACATCTGT 1457  
QY 1392 CGTCCCGGCGCTGT 1451  
DB 1458 GGGGTGAGGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1514  
QY 1452 TAGACCTGT 1511  
DB 1515 GGGGCTCAGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1574  
QY 1512 CCGTGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1571  
DB 1575 GCGATGT 1634  
QY 1572 TGT 1618  
DB 1635 CGT 1694  
QY 1619 --GGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1676  
DB 1695 CCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1754

QY 1677 GACCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1718  
DB 1755 GACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1796  
RESULT 15  
ADH71653  
ID ADH71653 standard; DNA; 1887 BP.  
XX  
AC ADH71653;  
XX  
DT 25-MAR-2004 (first entry)  
XX  
DE Human gene of the invention NOV22b SEQ ID NO:549.  
XX  
KW ds; gene; human; cytosolic; immunomodulator; neuroprotective; nootropic;  
KW anorectic; antidiabetic; antimicrobial; antipneumonic; gene therapy;  
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;  
KW obesity; diabetes; infectious disease; metabolic syndrome X;  
XX  
OS Homo sapiens.  
XX  
PN W02003102155-A2.  
XX  
PD 11-DEC-2003.  
XX  
PE 03-JUN-2003; 2003WO-US017430.  
XX  
PR 03-JUN-2002; 2002US-0385120P.  
PR 04-JUN-2002; 2002US-0385784P.  
PR 05-JUN-2002; 2002US-0386041P.  
PR 06-JUN-2002; 2002US-0386047P.  
PR 06-JUN-2002; 2002US-0386376P.  
PR 06-JUN-2002; 2002US-0386453P.  
PR 06-JUN-2002; 2002US-0386864P.  
PR 06-JUN-2002; 2002US-0387016P.  
PR 07-JUN-2002; 2002US-0387966P.  
PR 07-JUN-2002; 2002US-0388161P.  
PR 07-JUN-2002; 2002US-0388931P.  
PR 07-JUN-2002; 2002US-0388942P.  
PR 07-JUN-2002; 2002US-0389711P.  
PR 07-JUN-2002; 2002US-0389762P.  
PR 08-JUN-2002; 2002US-0389762P.  
PR 08-JUN-2002; 2002US-0389762P.  
PR 10-JUN-2002; 2002US-0387400P.  
PR 10-JUN-2002; 2002US-0387535P.  
PR 11-JUN-2002; 2002US-0387610P.  
PR 11-JUN-2002; 2002US-0387625P.  
PR 11-JUN-2002; 2002US-0387634P.  
PR 11-JUN-2002; 2002US-0387688P.  
PR 11-JUN-2002; 2002US-0387696P.  
PR 11-JUN-2002; 2002US-0387702P.  
PR 11-JUN-2002; 2002US-0387836P.  
PR 11-JUN-2002; 2002US-0387859P.  
PR 12-JUN-2002; 2002US-0387933P.  
PR 12-JUN-2002; 2002US-0387934P.  
PR 12-JUN-2002; 2002US-0387960P.  
PR 12-JUN-2002; 2002US-0388022P.  
PR 12-JUN-2002; 2002US-0388096P.  
PR 13-JUN-2002; 2002US-0389123P.  
PR 14-JUN-2002; 2002US-0389118P.  
PR 14-JUN-2002; 2002US-0389120P.  
PR 14-JUN-2002; 2002US-0389144P.  
PR 14-JUN-2002; 2002US-0389146P.  
PR 17-JUN-2002; 2002US-0389729P.  
PR 17-JUN-2002; 2002US-0389742P.  
PR 18-JUN-2002; 2002US-0389884P.  
PR 19-JUN-2002; 2002US-0390066P.  
PR 19-JUN-2002; 2002US-0390209P.  
PR 21-JUN-2002; 2002US-0390763P.  
PR 17-JUL-2002; 2002US-0396706P.  
PR 06-AUG-2002; 2002US-0401628P.  
PR 09-AUG-2002; 2002US-0402156P.



Db 1098 CATTGCGGCATATGAGCTGGCGAGGCACAGCTGCTGTGAGACTGACTGTGGGTCCCC 1157  
QY 1104 -----GCTGACCTTGCCCATAGTGGGAAACAGAGTGCAGAGGGGGCGCCC 1151  
Db 1158 ACCACCTCTCAAGCTAGCCACAGCACAGCTGTGACCCCGCGGGAACGGGATCTGA 1217  
QY 1152 CGGACCTCGGACATGCGCGCTCCGCTCGCACTGCTGCCAGAGGTGAGGGGACGCTGA 1211  
Db 1218 TGCTTCACCCCACTTCGCTGCTCTCTCTTCGCAAGGTGCGCACATGGGCCCC 1277  
QY 1212 GTCTGAGCCAGCCGTGACAGGAGTGAACCGCACCTCAGAGGCTGTGAGCTGGGG 1271  
Db 1278 TACCGACCGTGGAGTCAAGTACAGACAGGAGCACAGCTGCTTTGTCCAGTGGCC 1337  
QY 1272 TCCCGGCGGCAAGCCGACCCAGTGTGATGTTCCAAATCCAGTACAAACAGAGCCAGA 1331  
Db 1338 GGATCAGCGGCTATCCCGGATCCGATGTACCAAGATCCAGTACAGTACAGCTCGGCTGA 1397  
QY 1332 TGAGACCTCATCTACCGGATTTGTCCAGCTTCAGCCACCACTTCTGTGAGACACT 1391  
Db 1398 TGCATCTCTGTTACAGAGTATCCCGGAGAGCCGCTGTTCTGTGACGACCT 1457  
QY 1392 CGTCCCGGCGCTGACTATGACCTTGCCTGTGCTTGTACCGGCGCTGGGCTTC 1451  
Db 1458 GAGCTCAGGCGGACCTACAGATCTGTGCTGTGCGCTGTATGAGAC--AGCGCAC 1514  
QY 1452 TGACTTCAGGCGCACAGGCTGTGGGCTGTGCCATTTCTCAGCTGCGGCTCGCC 1511  
Db 1515 GGGGCTCAGGCGCACAGGCTGTGGGCTGTGGCGCTTCTCCACCGAACCTGGCTGCG 1574  
QY 1512 CCTGGCCAGGCGCTGACAGGCCCAAGTGTGGGCGGACCTGACCGTGGCGGTGGGG 1571  
Db 1575 GCCATGCGGGGCGCGCAGCTCCCTTCTGTGGCGGACGATGATCATCGGCTGGGCG 1634  
QY 1572 TGTGCTGCTGCTGCTTACTGTCTTCACTGTGGCTTGTGGTTC----- 1618  
Db 1635 CGTCATGCTAGCTCGGACTGCTTCTTCTGTGCTGCTAATGCGCTAACAGGTGCA 1694  
QY 1619 --GGGCTCGGGGGCGCGAAATGAGCGGCTTCCCTCAAGCTCAGCCAGTCCAGTCCA 1676  
Db 1695 CGGCGGCGAGCCCCCGGCAAGGCCAAGATTCGCGCGCTGTGCAAGCGTTTGTCTCCA 1754  
QY 1677 GACCAATGAGAGCCCGCACCCACACCCAGAGCCCAACCCGCC 1718  
Db 1755 GACCAACGCGGCTTGGGCCCCCAGCGCCCGCGCGCC 1796

Search completed: November 20, 2004, 08:43:10  
Job time : 1338 secs

*This Page Blank (uspto)*









QY 241 CGCATGTCATCAATCCCGATTGGGGCCCGCGCTTTGGGAGACTTCGAGAGCCTGGCTTCC 300  
 DB 370 CGCATGTCATCAATCCCGATTGGGGCCCGCGCTTTGGGAGACTTCGAGAGCCTGGCTTCC 429  
 QY 301 CTCCACTTTGACGGGACAGGCTGTGTGAGCTGGGACACGGGAGCCTCCGGGGCCCGGTC 360  
 DB 430 CTCCACTTTGACGGGACAGGCTGTGTGAGCTGGGACACGGGAGCCTCCGGGGCCCGGTC 489  
 QY 361 AATCTGACGACCTATCTCTACGGGCAACAGCTGGGGCCGATGCGCGCGGAGACCTTC 420  
 DB 490 AATCTGACGACCTATCTCTACGGGCAACAGCTGGGGCCGATGCGCGCGGAGACCTTC 549  
 QY 421 GACGACTTCTTAGAGAGCTGTGAGGACCTGAGACTGTCTCTCAACAACCTCCGGAGGTG 480  
 DB 550 GACGACTTCTTAGAGAGCTGTGAGGACCTGAGACTGTCTCTCAACAACCTCCGGAGGTG 609  
 QY 481 CCTGGGGCCGGGATCGGGCGCATGCTGTGCTGACACCTTCACACTTGAACCTTAACTTT 540  
 DB 610 CCTGGGGCCGGGATCGGGCGCATGCTGTGCTGACACCTTCACACTTGAACCTTAACTTT 669  
 QY 541 ATTGACGACCTGCCCCAGAGCCCTTGGGCGCAGCTCGGACCTCTCCCGCGGAGCTTC 600  
 DB 670 ATTGACGACCTGCCCCAGAGCCCTTGGGCGCAGCTCGGACCTCTCCCGCGGAGCTTC 729  
 QY 601 ACTTCAACCGGCTGGGACAGCTGTGCTCGGACCCGCTTTCTCTGTGGGCGTGAATGA 660  
 DB 730 ACTTCAACCGGCTGGGACAGCTGTGCTCGGACCCGCTTTCTCTGTGGGCGTGAATGA 789  
 QY 661 GAGGCTCTCCCGCCCGCTGTGTGCTGAGCTTTAGCGGGAACCTCCCTGCACTGCACTGT 720  
 DB 790 GAGGCTCTCCCGCCCGCTGTGTGCTGAGCTTTAGCGGGAACCTCCCTGCACTGCACTGT 849  
 QY 721 GAGCTGT 780  
 DB 850 GAGCTGT 909  
 QY 781 CCGGCGCTGGGCGGCGCTACTTCTGTGAGAGTGCCTCCGAGGCGAGTTCCTCTGTGAGCGG 840  
 DB 910 CCGGCGCTGGGCGGCGCTACTTCTGTGAGAGTGCCTCCGAGGCGAGTTCCTCTGTGAGCGG 969  
 QY 841 CCGCTCATTTGCGCGCAACAGCGAGCGCTTGTGGGTGTGTGAAGCCAGCGGGCCAGCGTG 900  
 DB 970 CCGCTCATTTGCGCGCAACAGCGAGCGCTTGTGGGTGTGTGAAGCCAGCGGGCCAGCGTG 1029  
 QY 901 CCGTGTCCGGGCGCTGT 960  
 DB 1030 CCGTGTCCGGGCGCTGT 1089  
 QY 961 TTGGTTGGCAACTCTCTCCGAGCCCGGGCTTTCCCAACGGGACCTTGAATTTGGGGTG 1020  
 DB 1090 TTGGTTGGCAACTCTCTCCGAGCCCGGGCTTTCCCAACGGGACCTTGAATTTGGGGTG 1149  
 QY 1021 ACCGGCGCTGGGAGCGTGTGGGCGTACACTGTGATGCGCACCAACCTGTGTGTGAGGCC 1080  
 DB 1150 ACCGGCGCTGGGAGCGTGTGGGCGTACACTGTGATGCGCACCAACCTGTGTGTGAGGCC 1209  
 QY 1081 ACAGCCCGGAGTGAACGTGGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140  
 DB 1210 ACAGCCCGGAGTGAACGTGGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1269  
 QY 1141 GGGGGGCGCGC----- 1149  
 DB 1270 GGGGGGCGCGC----- 1329  
 QY 1150 -----CCGGGGCTTGGACATGCGCGCTCCGCTGTGCACTGTGTGCGGAGGGTGAAGGAGCG 1206  
 DB 1330 CCGGCGCGGGCGCTTGGACATGCGCGCTCCGCTGTGCACTGTGTGCGGAGGGTGAAGGAGCG 1389  
 QY 1207 CTGAGAGCTTGAAGCAGGCGGTGTGAGGTGACGAGGAGTGAACCGGCACTTCAAGGCTGTGTAGC 1266  
 DB 1390 CTGAGAGCTTGAAGCAGGCGGTGTGAGGTGACGAGGAGTGAACCGGCACTTCAAGGCTGTGTAGC 1449  
 QY 1267 TGGGGTCCCGGGCGGCGAGCCGAGCCAGTGTGATGTTCCAATTCAGTACAGACGAGC 1326

DB 1450 TGGGGTCCCGGGCGAGCCGAGCCAGCCAGCGTGTGATGTTCCAATTCAGTACAGACGAGC 1509  
 QY 1327 GAAAGTGAACCTTCATTTACCGGATTTGTCCAGCTTCACGACCACTTCTGTGTAAG 1386  
 DB 1510 GAAAGTGAACCTTCATTTACCGGATTTGTCCAGCTTCACGACCACTTCTGTGTAAG 1569  
 QY 1387 CACTGTGTCCCGGCGCTGACTGATGACCTGTGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1446  
 DB 1570 CACTGTGTCCCGGCGCTGACTGATGACCTGTGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1629  
 QY 1447 CCTTGTGACTTCAAGGCGACACGAGT 1506  
 DB 1630 CCTTGTGACTTCAAGGCGACACGAGT 1689  
 QY 1507 TGGCCCGGTGTGCAAGGCGCTGTGAGGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1566  
 DB 1690 TGGCCCGGTGTGCAAGGCGCTGTGAGGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1749  
 QY 1567 GGGGGT 1626  
 DB 1750 GGGGGT 1809  
 QY 1627 GGGGGCGGAAATGGCGCGCTTCCCTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 1686  
 DB 1810 GGGGGCGGAAATGGCGCGCTTCCCTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 1869  
 QY 1687 GGGCCGAGCCCGACACCCAGGCGCCAGCCGCGCGGAGCCCGCGCGCGCGCGCGCGCGCGCG 1746  
 DB 1870 GGGCCGAGCCCGACACCCAGGCGCCAGCCGCGCGGAGCCCGCGCGCGCGCGCGCGCGCGCG 1929  
 QY 1747 AGCTGTCTTGTGAACTGTGGAATGTGCGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1806  
 DB 1930 AGCTGTCTTGTGAACTGTGGAATGTGCGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1989  
 QY 1807 GCTTGGGCGCGGAGGAGCACTGTGTGATGAGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1866  
 DB 1990 GCTTGGGCGCGGAGGAGCACTGTGTGATGAGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2049  
 QY 1867 GTAGGAG 1874  
 DB 2050 GTAGGAG 2057

RESULT 3  
 US-09-897-214-7  
 ; Sequence 7, Application US/09897214  
 ; Patent No. US20020076779A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Thayer, Edward C.  
 ; APPLICANT: Sheppard, Paul O.  
 ; APPLICANT: Presnell Scott R.  
 ; TITLE OF INVENTION: Leucine-Rich Repeat Proteins, Zlrr7,  
 ; TITLE OF INVENTION: Zlrr8, and Zlrr9  
 ; FILE REFERENCE: 01-27  
 ; CURRENT APPLICATION NUMBER: US/09/897,214  
 ; CURRENT FILING DATE: 2001-07-02  
 ; PRIOR APPLICATION NUMBER: 60/215,446  
 ; PRIOR FILING DATE: 2000-06-30  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 7  
 ; LENGTH: 1653  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)...(1653)  
 ; US-09-897-214-7  
 Query Match 85.4%; Score 1629; DB 9; Length 1653;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1651; Conservative 0; Indels 2; Gaps 2;

QY	1	ATGAGCCCCCGCGCTCTGCTGCTGCTGCTGCGCAGTGAAGCGCGCGCTTGCCCGTGGCC	60
Db	1	ATGGCCCGCCCGCTCTGCTGCTGCTGCTGCGCAGTGAAGCGCGCGCTTGCCCGTGGCC	60
QY	61	TGCGTGTGCAGAACCTGTGCCAGTGGCTCAGACGCTCTGTGGCCACCGAGGCGTGTG	120
Db	61	TGCGTGTGCAGAACCTGTGCCAGTGGCTCAGACGCTCTGTGGCCACCGAGGCGTGTG	120
QY	121	TTTGTGCCGCCCAAGCTGAGACCGCGCGCAGTGAAGCTTGCGGCTGGCTGACAACTTCATC	180
Db	121	TTTGTGCCGCCCAAGTGAACCGCGCGCAGTGAAGCTTGCGGCTGGCTGACAACTTCATC	180
QY	181	CAGGCGCTGGGGGGCCCCCTGACTTCCGCAACATGACGGGACTGTGTGACCTTGACCTGTCT	240
Db	181	CAGGCGCTGGGGGGCCCCCTGACTTCCGCAACATGACGGGACTGTGTGACCTTGACCTGTCT	240
QY	241	CGCAATGCGCATACCCCGATTGGGGCCCGCGCTTTGGGGGACCTGAGAGCTGCGGTTCC	300
Db	241	CGCAATGCGCATACCCCGATTGGGGCCCGCGCTTTGGGGGACCTGAGAGCTGCGGTTCC	300
QY	301	CTCCACCTTTAGCGCAACAGGCTGTGTGAAGCTGGGCACCGGAGCTTCGGGGCCCGCTC	360
Db	301	CTCCACCTTTAGCGCAACAGGCTGTGTGAAGCTGGGCACCGGAGCTTCGGGGCCCGCTC	360
QY	361	AATCGACGACACCTCATCTCCTGACGCGCAACAGCTGGGGCGGACTGGCGCGCGGAAACCTTC	420
Db	361	AATCGACGACACCTCATCTCCTGACGCGCAACAGCTGGGGCGGACTGGCGCGCGGAAACCTTC	420
QY	421	GACGACTTCTAGAGAGCTGTGAGGACCTTGAACCTGTCTTACAAACACTTCGGGAGGTG	480
Db	421	GACGACTTCTAGAGAGCTGTGAGGACCTTGAACCTGTCTTACAAACACTTCGGGAGGTG	480
QY	481	CCCTGGGGCGGACTGGGGCGCATGGCTGCCCTTGACAACCTTACACTTGAACATACTT	540
Db	481	CCCTGGGGCGGACTGGGGCGCATGGCTGCCCTTGACAACCTTACACTTGAACATACTT	540
QY	541	ATTGACGACATGCCCCCGAGCGGCTTTGGCCGACATCGGTACAGCTTCCGCGCTTGACCTC	600
Db	541	ATTGACGACATGCCCCCGAGCGGCTTTGGCCGACATCGGTACAGCTTCCGCGCTTGACCTC	600
QY	601	ACCTTCACAACCGCTTGCGCAACGTGGCTTCGGAACCGGCTTTTCTCTGTGGGCGTATGCA	660
Db	601	ACCTTCACAACCGCTTGCGCAACGTGGCTTCGGAACCGGCTTTTCTCTGTGGGCGTATGCA	660
QY	661	GAGGCGCTTCCCGCGCCCCCTGGTGTGAGCTTTAGCGGAAACCCCTGCACTGCAACGT	720
Db	661	GAGGCGCTTCCCGCGCCCCCTGGTGTGAGCTTTAGCGGAAACCCCTGCACTGCAACGT	720
QY	721	GAGCTGTGTGACTGCGGCGGCTGGCGCGGCGCGACGACTTGAAACGTTGCGCTCCCG	780
Db	721	GAGCTGTGTGACTGCGGCGGCTGGCGCGGCGCGACGACTTGAAACGTTGCGCTCCCG	780
QY	781	CCCGGCGCTGGCGCGGCGCTACTCTTGCGGCAATGCCCGAGGCGGAGTTTCTCTGTGAAGCG	840
Db	781	CCCGGCGCTGGCGCGGCGCTACTCTTGCGGCAATGCCCGAGGCGGAGTTTCTCTGTGAAGCG	840
QY	841	CCCTCATTTGCGCGCACACGACGACCTTGTGTGATGAGGCGACGCGGACACGCTG	900
Db	841	CCCTCATTTGCGCGCACACGACGACCTTGTGTGATGAGGCGACGCGGACACGCTG	900
QY	901	CGGTGCGGGGCGCTTGAGTGAACCCCGCGCTACATGCACTGGGTGCTGACGACCGG	960
Db	901	CGGTGCGGGGCGCTTGAGTGAACCCCGCGCTACATGCACTGGGTGCTGACGACCGG	960
QY	961	TTGTGTGCAACTCTTCCCGAGCGGGCTTTCCCAACGGGACCTTGAGATTGGGCTG	1020
Db	961	TTGTGTGCAACTCTTCCCGAGCGGGCTTTCCCAACGGGACCTTGAGATTGGGCTG	1020
QY	1021	ACCGGCGCTGGGAGAGCTGGGGGCTACACTTGACATGCGACCAACCTGTGTTGAGGCG	1080
Db	1021	ACCGGCGCTGGGAGAGCTGGGGGCTACACTTGACATGCGACCAACCTGTGTTGAGGCG	1080

QY	1081	ACAGCCGAGTAACTGCGGGGTCTGGCCCTTGGCCCATGGTGGGAACAGCAATGCGGAG	1140
Db	1081	ACAGCCGAGTAACTGCGGGGTCTGGCCCTTGGCCCATGGTGGGAACAGCAATGCGGAG	1140
QY	1141	GGGGGCCCGCCCGGGGCGCTCGGAATGCGCGCCCTCCGCTCGCATCTGCTCGCGAGGGGTAG	1200
Db	1141	GGGGGCCCGCCCGGGGCGCTCGGAATGCGCGCCCTCCGCTCGCATCTGCTCGCGAGGGGTAG	1200
QY	1201	GGGACGCTGAGCTTGAGCCACCGGTGAGGTGACGAGAGTGACCGCCCACTCAAGGCTG	1260
Db	1201	GGGACGCTGAGCTTGAGCCACCGGTGAGGTGACGAGAGTGACCGCCCACTCAAGGCTG	1260
QY	1261	GTGAGCTGAGGCTCCCGGGCGGCGCAGCCGACCCAGTGGAGATGTTCCAAATCAGTACAC	1320
Db	1261	GTGAGCTGAGGCTCCCGGGCGGCGCAGCCGACCCAGTGGAGATGTTCCAAATCAGTACAC	1320
QY	1321	AGCAGCGAAGATGAGACCTTCATCTACCGGATTGTGCCAGCTTCAGCCACCACTTCTG	1380
Db	1321	AGCAGCGAAGATGAGACCTTCATCTACCGGATTGTGCCAGCTTCAGCCACCACTTCTG	1380
QY	1381	CTGAAGACACCTTCGCTCCCGGGCTGACTATGACTCTTGCTGCTGCTGCTGTACCGGGCC	1440
Db	1381	CTGAAGACACCTTCGCTCCCGGGCTGACTATGACTCTTGCTGCTGCTGCTGTACCGGGCC	1440
QY	1441	GCTGGGCGCTTGACCTCACCGGACACAGGCTGCGGAGCTGTGGCCATTTCTCCAGCGTG	1500
Db	1441	GCTGGGCGCTTGACCTCACCGGACACAGGCTGCGGAGCTGTGGCCATTTCTCCAGCGTG	1500
QY	1501	CCGGCTTCGCGCCCTGTGTGCACGCGCTTGAAGGCCACAGCTGTGGGGAGGCCCTGACCGTG	1560
Db	1501	CCGGCTTCGCGCCCTGTGTGCACGCGCTTGAAGGCCACAGCTGTGGGGAGGCCCTGACCGTG	1560
QY	1561	GCCGTGGGGGGGTGTGCTGTGGTGGTGC-CCTTACTGGTCTTACAGTGG-CCTGCTGGTTC	1618
Db	1561	GCCGTGGGGGGGTGTGCTGTGGTGGTGC-CCTTACTGGTCTTACAGTGG-CCTGCTGGTTC	1620
QY	1619	GGGGCCCGGGGGCGGAAATGGCGGCTCTCCCC	1651
Db	1621	GGGGCCCGGGGGCGGAAATGGCGGCTCTCCCC	1653
RESULT 4			
US-10-093-463-189			
; Sequence 189, Application US/10093463			
; Publication No. US20030208039A1			
; GENERAL INFORMATION:			
; APPLICANT: Padigaru, Muralidhara			
; APPLICANT: Shenoy, Suresh			
; APPLICANT: Kekuda, Ramesh			
; APPLICANT: Gusev, Vladimir			
; APPLICANT: Pochart, Pascal			
; APPLICANT: Zhong, Mei			
; APPLICANT: Rastelli, Luca			
; APPLICANT: Mezes, Peter			
; APPLICANT: Smithson, Glenda			
; APPLICANT: Guo, Xiaojia			
; APPLICANT: Gerlach, Valerie			
; APPLICANT: Caeman, Stacie			
; APPLICANT: Boldog, Ferenc			
; APPLICANT: Li, Li			
; APPLICANT: Zephusen, Bryan			
; APPLICANT: Tchernev, Velizar			
; APPLICANT: Gangolli, Esna			
; APPLICANT: Vernet, Corine			
; APPLICANT: Pena, Carol			
; APPLICANT: Burgess, Catherine			
; APPLICANT: Liu, Xiaohong			
; APPLICANT: Spylek, Kimberly			
; APPLICANT: Gorman, Linda			
; APPLICANT: Spaderna, Steven			
; APPLICANT: Voss, Edward			
; APPLICANT: Malyankar, Uriel			
; APPLICANT: Anderson, David			



QY 1321 AGCAGGAGATGAGACCCCTCATCTACCGGATTTGCCAGCCTCCAGCCACTTCTCTG 1380  
| | | | |  
Db 1791 AGCAGGAGATGAGACCCCTCATCTACCGGATTTGCCAGCCTCCAGCCACTTCTCTG 1850  
QY 1381 CTGAACACCTCTGCTCCCGGCGCTGACTATGACCTTGTGCTGCTGCTTGTACCGGCGC 1440  
| | | | |  
Db 1851 CTGAACACACCCGCTCCCGGCGCTGACTATGACCTTGTGCTGCTGCTTGTACCGGCGC 1910  
QY 1441 GCTGGGCGCTCTGACTCATCGGCGACAGGCTGTGGGCTGTGGCCATTTCACAGCTG 1500  
| | | | |  
Db 1911 GCTGGGCGCTCTGACTCATCGGCGACAGGCTGTGGGCTGTGGCCATTTCACAGCTG 1970  
QY 1501 CGGCGCTGCGCCCTGTGGCCAGCCCTGTGAGGCGCCACAGTCTGTGGGCGGACCTGACCGTG 1560  
| | | | |  
Db 1971 CGGCGCTGCGCCCTGTGGCCAGCCCTGTGAGGCGCCACAGTCTGTGGGCGGACCTGACCGTG 2030  
QY 1561 GCCGTGGGCGGTGTGCTGTGTGCTGCTG 1586  
| | | | |  
Db 2031 GCCGTGGGCGGTGTGCTGTGTGCTGCTG 2056

## RESULT 5

US-09-768-826-21  
; Sequence 21, Application us/09768826  
; Patent No. US20020012966A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi et al.  
; TITLE OF INVENTION: 18 human secreted proteins  
; FILE REFERENCE: P512P1  
; CURRENT APPLICATION NUMBER: US/09/768,826  
; CURRENT FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: PCT/US00/22350  
; PRIOR FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: 60/148,759  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 21  
; LENGTH: 1872  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-768-826-21

Query Match 81.9%; Score 1563.4; DB 9; Length 1872;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1564; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 344 GCTCCCGGCGCGCTCAATCTGACGACCTCATCTGAGGCGAACAGAGCTGGGCGCA 403  
| | | | |  
Db 1 GCTCCCGGCGCGCTCAATCTGACGACCTCATCTGAGGCGAACAGAGCTGGGCGCA 60  
QY 404 TCGCGCGCGGAGCCTTTCAGCACTTCTTGAAGAGCCTGAGGACCTGTGACTTCTTACA 463  
| | | | |  
Db 61 TCGCGCGCGGAGCCTTTCAGCACTTCTTGAAGAGCCTGAGGACCTGTGACTTCTTACA 120  
QY 464 ACAACCTTCGCGAGGTGCTGCGGCGGCAATCGGCGCATGCTGCTGCAACACCTTCA 523  
| | | | |  
Db 121 ACAACCTTCGCGAGGTGCTGCGGCGGCAATCGGCGCATGCTGCTGCAACACCTTCA 180  
QY 524 ACCTGACCATTAACCTTATGAGCGACGTGCGGCGGCGCTTGGCCGAGTGTGCTGAGC 583  
| | | | |  
Db 181 ACCTGACCATTAACCTTATGAGCGACGTGCGGCGGCGCTTGGCCGAGTGTGCTGAGC 240  
QY 584 TCTCCCGCTGAGACTTCACTTCAACCGGCTGTGCGACGCTGCTGCGACCGGCTTTTCT 643  
| | | | |  
Db 241 TCTCCCGCTGAGACTTCACTTCAACCGGCTGTGCGACGCTGCTGCGACCGGCTTTTCT 300  
QY 644 CTCTGGGCGCTGATGACAGAGGCTTCTCCGCGCCCTGTGTCTTGAAGCTTTTGAAGGAAAC 703  
| | | | |  
Db 301 CTCTGGGCGCTGATGACAGAGGCTTCTCCGCGCCCTGTGTCTTGAAGCTTTTGAAGGAAAC 360  
QY 704 CCTGCACTGCACTGTGAGCTGTGTGCTGTGGCGGCGCTGTGGCGGCGGACGACCTGTG 763

Db 361 CCTCACTGCAAACTGTGAGCTGTGTGCTGTGCGCGGCTGTGCGCGGCGGACGACTGTG 420  
QY 764 AAACGTGGCGCTCCCGCGCGGCTGTGCGCGGCTGTGCGCGGCTGTGCGCGGCTGTGCGCGG 823  
| | | | |  
Db 421 AAACGTGGCGCTCCCGCGCGGCTGTGCGCGGCTGTGCGCGGCTGTGCGCGGCTGTGCGCGG 480  
QY 824 AGTTCTCTGTGAGCGCGGCTCATTTGCTCCGACACGACGAGCCTGTGGGTGTGAGGAG 883  
| | | | |  
Db 481 AGTTCTCTGTGAGCGCGGCTCATTTGCTCCGACACGAGGAGCCTGTGGGTGTGAGGAG 540  
QY 884 GCCACGCGGCGCACGCTGTGTGTGCGGCGCTGTGTGTGACCCGCGCTTACATGCACTGTGG 943  
| | | | |  
Db 541 GCCACGCGGCGCACGCTGTGTGTGCGGCGCTGTGTGTGACCCGCGCTTACATGCACTGTGG 600  
QY 944 TCGGCTGCGAGACCGGTTGTGTGTGCACTCTCCCGAGCCCGGCGCTTCCCGAACGGA 1003  
| | | | |  
Db 601 TCGGCTGCGAGACCGGTTGTGTGTGCACTCTCCCGAGCCCGGCGCTTCCCGAACGGA 660  
QY 1004 CTTTGAAGATTGGGTGACCGGCGCTGTGGGAGCGCTGTGGGCGCTTACACTGCACTGCGACCA 1063  
| | | | |  
Db 661 CTTTGAAGATTGGGTGACCGGCGCTGTGGGAGCGCTGTGGGCGCTTACACTGCACTGCGACCA 720  
QY 1064 ACCCTGTGTGTGAGGCGCACAGCCGAGTAACTGTGCGGTGTGTGCTTGTCCCATGTGTG 1123  
| | | | |  
Db 721 ACCCTGTGTGTGAGGCGCACAGCCGAGTAACTGTGCGGTGTGTGCTTGTCCCATGTGTG 780  
QY 1124 GAAACAGAGTGGCGGAGGCGGCGCGCGCGGCGCTGTGGACATGCGCGGCTCCGCTGCCA 1183  
| | | | |  
Db 781 GAAACAGAGTGGCGGAGGCGGCGCGCGCGGCGCTGTGGACATGCGCGGCTCCGCTGCCA 840  
QY 1184 CTGCTGCGGAGGTGTGAGGCGGAGCGCTGTGAGTGTGAGCCAGCGGTGTGACGAGGTGA 1243  
| | | | |  
Db 841 CTGCTGCGGAGGTGTGAGGCGGAGCGCTGTGAGTGTGAGCCAGCGGTGTGAGGAGGTGA 900  
QY 1244 CCGGCACTTCAAGGCTGTGTGAGCTGTGGGTTCCCGGCGGCGCACGCGGACCTGATGTGT 1303  
| | | | |  
Db 901 CCGGCACTTCAAGGCTGTGTGAGCTGTGGGTTCCCGGCGGCGCACGCGGACCTGATGTGT 960  
QY 1304 TCCAAATCCAGTAAACAAGACGAGCGAAGTGAACCTCATCTACCGGATTTGCCAGGCT 1363  
| | | | |  
Db 961 TCCAAATCCAGTAAACAAGACGAGCGAAGTGAACCTCATCTACCGGATTTGCCAGGCT 1020  
QY 1364 CCAGCACAACCTTCTGTGTGAAGACCTGTGTCCCGGCGGTGACTATGACCTTGTGCTGCG 1423  
| | | | |  
Db 1021 CCAGCACAACCTTCTGTGTGAAGACCTGTGTCCCGGCGGTGACTATGACCTTGTGCTGCG 1080  
QY 1424 TGGCTTGTACCGGCGGCTGTGGCCCTTGTGACTTCACTGACGCGGACCAAGCTGTGCTGTG 1483  
| | | | |  
Db 1081 TGGCTTGTACCGGCGGCTGTGGCCCTTGTGACTTCACTGACGCGGACCAAGCTGTGCTGTG 1140  
QY 1484 CCCATTCTCCAGCTGTGCGGCGCTGCGCCCTGTGTGCAAGCGGCTGTGAGGCGGACGAGCTGTG 1543  
| | | | |  
Db 1141 CCCATTCTCCAGCTGTGCGGCGCTGCGCCCTGTGTGCAAGCGGCTGTGAGGCGGACGAGCTGTG 1200  
QY 1544 GCGGAGCCCTTGAACCGTGTGCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1603  
| | | | |  
Db 1201 GCGGAGCCCTTGAACCGTGTGCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1260  
QY 1604 TGGCTTG 1663  
| | | | |  
Db 1261 TGGCTTG 1320  
QY 1664 ACCTGACAGTCCCAAGCAATGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1723  
| | | | |  
Db 1321 ACCTGACAGTCCCAAGCAATGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380  
QY 1724 GCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1783  
| | | | |  
Db 1381 GCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440  
QY 1784 GTTATGCAAGCGGCTGTGGAGAGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1843  
| | | | |











QY 1677 GACCAATGAGGCCCGACCCCAACACCAAGGCCACCCGCC 1718  
Db 1755 GACCAAGGCGCCCTGCGGCCCGCCACGCCACGCCCGCCGCC 1796

RESULT 10  
US-10-221-097-19  
; Sequence 19, Application US/10221097  
; Publication No. US2003014476A1  
; GENERAL INFORMATION:  
; APPLICANT: Agarwal, Pankaj  
; APPLICANT: Murdock, Paul R.  
; APPLICANT: Rizvi, Safia K.  
; APPLICANT: Smith, Randall F.  
; APPLICANT: Xiang, Zhaoxing  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; FILE REFERENCE: GP50016  
; CURRENT APPLICATION NUMBER: US/10/221,097  
; CURRENT FILING DATE: 2002-09-06  
; PRIOR APPLICATION NUMBER: PCT/US01/07143  
; PRIOR FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: 60/187,107  
; PRIOR FILING DATE: 2000-03-06  
; PRIOR APPLICATION NUMBER: 60/236,874  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/188,916  
; PRIOR FILING DATE: 2000-03-13  
; PRIOR APPLICATION NUMBER: 60/237,846  
; PRIOR FILING DATE: 2000-10-03  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 19  
; LENGTH: 1887  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-221-097-19

Query Match 32.7%; Score 624.8; DB 15; Length 1887;  
Best Local Similarity 62.4%; Pred. No. 1e-137;  
Matches 1074; Conservative 0; Mismatches 597; Indels 51; Gaps 4;

QY 45 GCGCTGCCCGCTGCGCTGCTGCGCAACCTGTCGAGTGCCTGACGACCTCTGTGC 104  
Db 78 CCCATGTCGCCCGCTGCGCTGCGCTGCGCAACCTGTCGAGTGCCTGTCGTGCC 137

QY 105 CCAACCGAGGCGCTGCTGCTGCTGCTGCGCGCAACCTGTCGAGTGCCTGTCGTGCC 164  
Db 138 AGGGGCGAGGCGCTGCTGCTGCTGCGCAACCTGTCGAGTGCCTGTCGTGCC 197

QY 165 GCGTGAACAATTCTCAACGAGGCGCTGCGGCGCGCTGACCTTCCGCAACATGACGGGACTGT 224  
Db 198 GCGGAGCAACTTCAATCGCTGCTGCGCGCGCGCGGAGCTGCGCAACATGACGGGACTGT 257

QY 225 GGAACCTGACATGCTGTCGCAATGCAATGCAATGCGCAATGCGGCGCGCGCTTGGGAGCT 284  
Db 258 GCACTGAGGCTGTCGCGGGAACACATGCGCAATGCGGCGCGCGCTTGGGAGCT 317

QY 285 CGAAGCTGCTGCTGCTGCTGCGCGCAACGAGTGTGAGAGCTGCGGCGCGAG 344  
Db 318 GCGGCGCGCTGCTGCTGCGCGCAACGAGTGTGAGAGCTGCGGCGCGAG 377

QY 345 CCGTCCGCGGCGCGCTGCAATGCAACGCTGCTGCAACGCGCAACGAGCTGCGCGCAT 404  
Db 378 GCTGCGCGCGCTGCAATGCGGCACTGCTGCAACGAGCAACGAGCTGCGCGCAT 437

QY 405 CCGCGCGGAGGCTTTCGACGATTCCTAGAGAGCTGAGAGAGCTGCACTGCTCTTACAA 464  
Db 438 GCGCGCGGCGCGCTGATGATGATGCGGAGAGAGCTGAGAGAGCTGCACTCTCTTACAA 497

QY 465 CAACTCCGAGAGGAGCTGCGGCGCGGAGATGCGGCGCAATGCTGCGCGCAACCTCAA 524  
Db 498 CAACTCCGAGAGGAGCTGCGGCGCGGAGCTGCGGCGCGCTGCGGCAACGCTCAACGCTTGG 557

QY 525 CTTGACCAATTAACCTTATTAGACGACATGCGCCCGGCGCGCTTGGCCAGCTCGATCACT 584  
Db 558 CTTGACCAATTAACCTTATTAGACGACATGCGCCCGGCGCGCTTGGCCAGCTCGATCACT 617

QY 585 CTTCCGCGCTGAGACCTTCACTCCAGCGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCT 644  
Db 618 GCGCGCGCTGAGACCTTCACTCCAGCGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCT 677

QY 645 TCGTGGGCG-----GTGATGAGAGGCGCTGCGCGCGCGCGCGCTGCTGAGCTT 692  
Db 678 CCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTGAGCTT 737

QY 693 TAGCGGGAACCCCTGCGACCTGCACTGAGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCG 752  
Db 738 TGGCGGGAACCCCTGCGACCTGCACTGAGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCG 797

QY 753 GAGGACCTTGGAAACGTCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 812  
Db 798 GAGGACCTTGGAAACGTCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 857

QY 813 GCGCGAGGCGGAGCTTCTCTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 872  
Db 858 GCGCGAGGCGGAGCTTCTCTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 917

QY 873 GGTGCTGGAAGGCGGAGCG 932  
Db 918 TGTGCGCGGAGCTGCG 977

QY 933 CATGCACTGGGTGCGCTGCTGCAACGCGGTGCTGCACTGCTGCGCGCGCGCGCGCGCGCTT 992  
Db 978 TGTGCGGTGCGGTGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 1037

QY 993 CCGCAACGAGACCTTGAAGATTGGGAGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 1052  
Db 1038 CCGCAACGAGACCTTGAAGATTGGGAGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 1097

QY 1053 CATGCGCAACCGCTGCTGCTGAGGCTGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 1103  
Db 1098 CATGCGCGCAACCGCTGCTGAGGCTGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 1157

QY 1104 -----GTTGGCGCTTGGCG 1151  
Db 1158 ACCACTCTTCAAGTGAACCAACGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1217

QY 1152 CCGGCGCGCTGCGAGCATGCGCGCGCTGCGCTGCGCACTGCTGCGAGGAGTGGAGGAGCGTGA 1211  
Db 1218 TGTCTTCAACCGCAACCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1277

QY 1212 GTCTGAGCGCGGCTGCGAGTGAACGAGTGAACCGCGCACTGAGGCTGCTGAGCTGAGG 1271  
Db 1278 TACGACCGTGGCGTCCAGGCTGAGTGAACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1337

QY 1272 TCCCGGCG 1331  
Db 1338 GAGTCAAGGCGCTATCCCGGCG 1397

QY 1332 TGAAGACCTGATTAACCGGATTGTCGAGCTTCCAGCGCACTTCTGCTGAAGCACT 1391  
Db 1398 TGAATCTGCTGATTAACGAGATGATCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1457

QY 1392 CGTCCCGCGCGCTGATTAACGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1451  
Db 1458 GCGCTGAGGCG 1514

QY 1452 TGACCTCAACGCGCGCAACGAGCTGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1511  
Db 1515 GCGGCTCAACGCGCGCAACGAGCTGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1574

QY 1512 CTTGTCACGCGCGCTGAGGCG 1571  
Db 1575 GCGATGCGGCG 1634

QY 1572 TGTGCTGTGGCTGCTTACTGTCTTCACTGTGCTTGTGTTTC----- 1618  
 Db 1635 CGCATCTGTAGCTCGGTACTGTCTTCACTTTCGTGTGCTAATCGGTACAGAGTGCA 1694  
 QY 1619 --GGGCGCGGGGGGCGGAATGGCGGCTCCCGCTCAAGCTCAGCGACGTCGAGCCCA 1676  
 Db 1695 CGCGGGCCAGCCCCCGGGGCAAGGCAAGATTCGCGGCGCTTTAGAGCGTTTGCTCCA 1754  
 QY 1677 GACCAATGAGGCGCCGACCCCAACCCAAAGGCGCCACCGCC 1718  
 Db 1755 GACCAAGCGCGCCCTGGGCCCAAGCCCAAGCCCGCGCCG 1796

## RESULT 11

US-09-815-626-1  
 ; Sequence 1, Application US/09815626  
 ; Patent No. US20020076752A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Glucksmann, Maria Alexandra  
 ; TITLE OF INVENTION: 3395, A NOVEL HUMAN LEUCINE-RICH REPEAT  
 ; FILE REFERENCE: 10448-031001  
 ; CURRENT APPLICATION NUMBER: US/09/815,626  
 ; PRIOR FILING DATE: 2001-03-23  
 ; PRIOR APPLICATION NUMBER: US 60/191,863  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2558  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (435) ... (2318)  
 US-09-815-626-1

Query Match 32.7%; Score 624.8; DB 9; Length 2558;  
 Best Local Similarity 62.4%; Pred. No. 9.9e-138;  
 Matches 1074; Conservative 0; Mismatches 597; Indels 51; Gaps 4;

QY 45 CGCGTGGCGGCTGCTGCTGCGGAGCACTGTCCGATGCTCAGACCTCTGTGCT 104  
 Db 512 CCATGTCCTCCCGCGCTGCGCTGCGGACACAGTGTGCTGCTGCTGCTGCTGCTGCT 571  
 QY 105 CCACGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 164  
 Db 572 AGGGGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 631  
 QY 165 GCGTGAACAATTCAATCAGGCGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 224  
 Db 632 GGCAGACAACTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 691  
 QY 225 GGAACCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 284  
 Db 692 GATCTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 751  
 QY 285 CGAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 344  
 Db 752 GCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 811  
 QY 345 CTTGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404  
 Db 812 GCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 871  
 QY 405 GCGCGCGGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 464  
 Db 872 GCGCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 931  
 QY 465 CAACCTCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 524  
 Db 932 CAACCTCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 991

QY 525 CCTGAGCAATACTTATGAGCACTGCTCCCGAGGCGCTTGGCCAGCTGAGTACGT 584  
 Db 992 CTTGCAACCAACAACCTGTGGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1051  
 QY 585 CTTCCGCGTGAACCTTCACTTCAACCGCTGAGCGCAAGCTGTGCTGCTGCTGCTGCT 644  
 Db 1052 GCGCGCGTGAACCACTTCAACCGCTGAGCGCAACATCCCAACCGACCCCACTTCTCTC 1111  
 QY 645 TGTGGGC-----GTGATGACAGAGCTCTCCCGCGCTGCTGCTGCTGCTGCTGCT 692  
 Db 1112 CGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1171  
 QY 693 TAGCGGAAACCCCTGCACTGCACTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 752  
 Db 1172 TGGCGGAAACCCCTGCACTGCACTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1231  
 QY 753 GGAAGCACTGGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 812  
 Db 1232 GAGCACTTGAAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1291  
 QY 813 GCCGAGGCGAGTCTCTCTGTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 872  
 Db 1292 GGGCGAGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1351  
 QY 873 GGTGCTGGAAGGCGCAGCGGCGCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 932  
 Db 1352 TGTGCTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1411  
 QY 933 CATGCACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 992  
 Db 1412 TGTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1471  
 QY 993 CCCCAAGGAACTTAGAGATTGGGATGACCGCGCTGAGGAGAGCTGAGGAGGCTGCTGCT 1052  
 Db 1472 CCCCAATGGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1531  
 QY 1053 CATGCGCACCAACCTGCTGTGAGGCGCAAGCGCCGAGTAGAATGCTGCTGCTGCTGCT 1103  
 Db 1532 CATTCGCGCAATGAGCTGTGAGGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1591  
 QY 1104 -----GCTGGGCTTGGCCCATGTGTGGAACAGCAATGCCGAGAGGCGCGCC 1151  
 Db 1592 ACCACCTCTCAAGCTGAGCAACAGCAAGCTGTGACCCCGCGGAGAGGAGATCTGA 1651  
 QY 1152 CGGCGCTTGTGACATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211  
 Db 1652 TGCTGCAACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1711  
 QY 1212 GTCTGAGCAGCGCTGTGAGGTGACGAGGTGACCGCACCTTGAAGGCTGTGAGTGTGAG 1271  
 Db 1712 TACCGACGTTGCTGTGAGGTGACGAGGTGACCGAGGCGCACAGCTGCTTGTTCAGTGGCC 1771  
 QY 1272 TCCGCGGCGGCGAGCGCAACCGAGTGTGATGTGCAATCCAGTCAACAGAGGAGAA 1331  
 Db 1772 GATCAGGCGCTTATCCCGGAGATCCGATGTGACAGATCCAGTCAACAGTCCGCTGA 1831  
 QY 1332 TGAGACCTTCATTCACCGGATTTGCTCCAGCTTCCAGCAACCACTTCTGCTGAGACCT 1391  
 Db 1832 TGACATTCCTGCTTACAGGATATCCCGGCGGAGAGCGCTGCTTCTGCTGACGAGCT 1891  
 QY 1392 CGTCCCGCGGCTGTATGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1451  
 Db 1892 GCGCTGAGGCGCGAAGCTTGTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1948  
 QY 1452 TGACCTCAAGGCGCAACAGGCTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511  
 Db 1949 GGGGCTTCAAGGCGCAACAGGCTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2008  
 QY 1512 CTTGTGCAACGCGCTTGAAGGCGCAAGTGTGAGGCGGAGCCCTGACCGTGTGCGGAG 1571  
 Db 2009 GCGATGCGGAGGCGCGAGCTCCTTCTGTGAGGCGAGATGATCATGCGCTGTGCGG 2068  
 QY 1572 TGTGCTGAGTGTGCTTACTGTGCTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1618











Sat Nov 20 18:13:49 2004

us-10-071-879-9.rnpb

Page 18

Job time : 1371 secs

---

[illegible]

165 GAGTGAACATTTATTCAGAGCCCTTGGGCCCCCTTGAATTCGGCAATGACGGGACTGCT 224  
Db GGCAACAACTTTCATGCGCTCCGTGCGCGCGCACTGGCAATGACAGGCGCTCT 1443  
225 GAGACCTGACATGTCGCAATGCCATGACCCGCAATGGGCGCGCGCTTGGGAGACT 284  
Db GCATCTGAGCTTGTGCGGAAACAACTCGGCACCTGCTGCGCGCTTGGCGAGCT 1503  
285 CGAGAGCTGCGTTCCTTCCATTCGAGACCTGATCTTGAAGGAGAGCTGGAGCGGAG 344  
Db GGGGCGCGCTGCGCTGCACTGGATGCAACCGGCTGACCTGATGAGGAGGCA 1563  
345 CTTCCGGGGCCCCCTTCAATCTGAGACCTGATCTTGAAGGAGAGAGCTGGCGAGT 404  
Db GTTGGCGCGCTGCTGCACTTGGCGCACTGATCTGAGCAACAACTGAGCGGAGCT 1623  
405 CGCGCGGAGGCTTTCAGACCTTCTAGAGAGCTGAGAGAGCTGAGAGCTGCTTACAA 464  
Db GGGGCGCGCGCGCTTGGATGATTTGTGCGAGACACTGAGAGAGCTGAGCTTCTTACA 1683  
465 CAACCTCCGAGAGTGCCTTGGGCGGCAATCGGCGCATGCTGCTGCGCAACCTTCAA 524  
Db CAACCTCCGAGAGTGCCTTGGGCGGCGCTTGGGCGCAACCTGAGAGAGCTTGGG 1743  
525 CTTGAGACCTTATTTAGAGCACTGCGCGCGCGCGCTTGGCGAGCTGGTACGT 584  
Db CTTGAGACCAACTGCTGCTTCTGTGCGCGCGCGCTTGGCGAGCTTGGCGAGCT 1803  
585 CTCCGCGCTGAGACTCAGCTCCCAACCGGCTGGGCAAGCTGAGCGGAGCGCTTTC 644  
Db GGGCGCGCTGAGACTCAGCTCCCAACCGGCTGGGCAAGCTGAGCGGAGCGCTTTC 1863  
645 TGTGGGC-----GTGATGCAAGGCTTCTCCGCGCGCTGCTGCTGAGCTT 692  
Db CCGCTGCGCTGCTGCGAGGCGCGCGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTT 1923  
693 TAGCGGGAACCCGCTGCACTGCACTGAGAGCTGAGCTGAGCTGAGCGGCGGCGGCG 752  
Db TAGCGGGAACCCGCTGCACTGCACTGAGAGCTGAGCTGAGCTGAGCGGCGGCGG 1983  
753 GAGCAGCTGGAACCTGCGGCTCCCGCGCGCGCTGAGCGGCGGCTTCTTGGAGCT 812  
Db GAGCAGCTGAGAGCTGCGGCTCCCGCACTGCTGAGCGGCGGCTTCTTGGAGCT 2043  
813 GCGCGAGGCGAGTCTCTGTGAGCGCGCTCATGCTGCGCGCAGCAGCGAGGCTTCTG 872  
Db GCGCGAGGAGAGTCTCTGTGAGCGCGCTCATGCTGCGCGCAGCAGCGAGGCTTCTG 2103  
873 GGTGTGGAAGGCGAGCGGCGCACTGAGCTGAGCGGCGCGCGGCGGCGGCGGCTTAC 932  
Db TGTGCGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2163  
933 CATGCACTGAGCTGCTGAGCAGCGGCTTGTGAGCACTCTCCGAGCGCGGCGCTT 992  
Db TGTGCGTGTGAGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGCTT 2223  
993 CCGCAACGGGACTTGAAGTTGGGCTGAGCGGCGGCTGAGGAGCGGAGGAGCTGAGCTG 1052  
Db CCGCAATGGGAGCTGAGCTGAGCTGAGCTGAGCGGCGGAGGAGGAGCTGAGCTGAG 2283  
1053 CATGCGCACCAACCTGAGTGTGAGGCGCAGCGCGAGTAGAACTGCGGCT----- 1103  
Db CATTTGGCGCAATGAGCTGAGGCGAGGCGAGCTGCTGTGAGCTGAGCTGAGCTGAG 2343  
2284 CATTTGGCGCAATGAGCTGAGGCGAGGCGAGCTGCTGTGAGCTGAGCTGAGCTGAG 2403  
1104 -----GCTGCGCTTGGCGCAATGAGTGTGAGGAGCAGGAGTCCGAGGCGGCGGCG 1151  
Db AACACTCTCTGAGCTGAGCAGCAAGCAAGCTGTGAGCGCGCGGAGCGGAGTCTGAG 2403  
1152 CGGCGCTTGGAGATGCGGCTTCCGCTGAGCTGCTGCGGAGGAGGAGGAGGAGCTGAG 1211  
Db TGTCTCACCGCACTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2463

1212 GTCTGAGCAGCCGCTGAGGAGTGAAGGAGTGAACCGGCACTTCAAGGCTGTGAGCTGAGG 1271  
Db TACGACCTGTGAGCTGAGGAGTGAAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGG 2523  
1272 TCCCGGCGCGCGCGAGCGGAGCGGAGTGTGAGTGTTCAGATTCAGATTCAGAGGAGGA 1331  
Db GAGTCAAGCGGCTTATCCGCGGAGTCCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2583  
1332 TGAGACCTCATCTTACCGGATTTGTGCGAGCTTCCAGGCAACCACTTCCGCTGAGACACT 1391  
Db TGACATCTCTGTACAGATGATTCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2643  
1392 GCTCCCGCGCGCTGAGCTTGAAGCTTGTGAGCTGAGCTTGTGAGCTGAGCTGAGCTG 1451  
Db GAGCTCAGGCGGAGCTTGAAGCTTGTGAGCTGAGCTTGTGAGCTGAGCTGAGCTGAG 2700  
1452 TGACCTCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1511  
Db GAGGCTCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2760  
1512 CTTGCGCAGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1571  
Db GCGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2820  
1572 TGTGCTGAGCTGAGCTTACTGAGTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTT 1618  
Db CTTGCTGAGCTGAGCTTACTGAGTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTT 2880  
1619 --GGGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1676  
Db GGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2940  
1677 GACCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1718  
Db GACCAAGCGGCTTGGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2982

RESULT 2  
US-09-620-312D-92  
Sequence 92, Application US/09620312D  
Patent No. 6569662  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Feiyun  
APPLICANT: Chen, Rui-hong  
APPLICANT: Zhao, Qing A.  
APPLICANT: Wehrman, Tom  
APPLICANT: Xue, Aidong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yundong  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: John Tillinghast  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and  
FILE REFERENCE: Polypeptides  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: pl\_Fl\_genes Version 1.0  
SEQ ID NO 92  
LENGTH: 2818  
TYPE: DNA



```

; TYPE: DNA
; ORGANISM: Homo sapiens
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2019)
US-09-063-950-3

```

Query Match	5.1%;	Score 98.2;	DB 3;	Length 2019;
Best Local Similarity	52.0%;	Pred. No. 5e-11;		
Matches 270; Conservative	0;	Mismatches 243;	Indels 6;	Gaps 2;

QY	1	CCGCCCCCTCTCTGCTGTGCTGTGCGCAGATGGAGGGGCGCGCTGCGCTGCTGGCTGATC	66
Dp	22	CTGCTGCGCTGCTCTGTCTACTGTGGCCCTGGGGGCTGGGGGTGACAGGGCTGCCATCCGGC	81
QY	67	TGCGAGAACTCTTCGGAGTGGCTTGACAACTCTGTGTGCCACCGAGGCTGTGTTTGTG	126
Dp	82	TGCGAATGTG---AGCCAGCCACAGCAAGATCTTGTGCACTGCCCCCAGGGGACCAAGGTTG	138
QY	127	CCGCCCCAAGTGTGAGACGGGCGGCACAGTGGAGCTGGGGGCTGGGCTGACAACTTCATCCAGGCC	186
Dp	139	CCCCAGAGATGTGCACCCGACACGGTGGGGCTGTACGTCTTTGAGAAAGGCAATCACATG	198
QY	187	CTGGGGGCCCTGACTTTCGCAACATGACGGGACTGTGTGACCTGACACTGTCTCGCAAT	246
Dp	199	CTCGAGCGCAGGAGCTTTGTGCGCGCTGCGGGGCTGCAGCTCTGGACCTGTCAAGAAAC	258
QY	247	GCCATCACCGGCAATTGGGGGCCGGGCGCTTTGGGGACCTCGAGAGACGTCGGTTCCTCCAC	306
Dp	259	CAGATGCCAGGCTGTGCCACAGCGGGGTCTTCCAGCACCCTGGCAACTTCAGAACCTTGGAC	318
QY	307	CTTGAACGGCAACAGGCTGTGTGAGCTGGGACCGGAGCCTTCGGGGGCCCTCGTCAATCTTG	366
Dp	319	CTGACGGCCAAAGGCTGTGATGAATTCACCAATGAGACTTTCGGTGGCTTGCGGCGCTTC	378
QY	367	CAGCACTCATCTTCAGCGGGCAACAGTGGGGCGGCATCGGCGGGGAGGCTTCAGAGAC	426
Dp	379	GAGGCGCTTACTTGGGCAAGAACCGCATTCGGCACATCCAGCTGTGTGCTT---CGAC	435
QY	427	TTCTTAGAGAGCTGGAGACTTGACCTGTCTTCAACAACTTCGCGCAGAGTGGCTCTGG	486
Dp	436	AGGCTGAGCCGCTCTGTGAGCTTCAAGCTGCAGAGCAACGAGTCTGGGGCACTGGCCCCG	495
QY	487	GCGGCAATGGGGCAATGCTGCGCTCTGCACACCTCAAC	525
Dp	496	CTGCGCTCTCCGCGCTGTGCTGCTGGAAGCTTCAGCCAC	534

RESULT 4  
US-09-063-950-1  
; Sequence 1, Application US/09063950C

```

; APPLICANT Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL LRSg PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MEI-019
; CURRENT APPLICATION NUMBER: US/09/063,950C
; CURRENT FILING DATE: 1998-04-21
; NUMBER OF SEQ. ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ. ID NO 1
; LENGTH: 2852
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (160)..(2178)
; US-09-063-950-1

```

Query Match	5.1%;	Score 98.2;	DB 3;	Length 2852;
Best Local Similarity	52.0%;	Pred. No. 5.3e-11;		
Matches 270;	Conservative 0;	Mismatches 243;	Indels 6;	Gaps 2;

QY	7	CCGCGCCTCTTCTCTGCTGTGCTGGCCAGTGAAGAGCGCGCTTCCCGCTGCCTGTGTC	66
Db	181	CTGCTGCCCTGCTCTCTGTACTGCGCTTGGGGGCTGGGGGTGCAAGGGCTGCCATCCGGC	240
QY	67	TGCGAAGACTGTGCCAGTGTGTGACACCTCTGTGTCCACCGAGGCTGTGCTTTGTG	126
Db	241	TGCGAGTGTG---AGCGAGCGACAGACAGTCTTCTGTGCACTGCCCCCAGGGAGCCACAGTG	297
QY	127	CCGCCCAACGTGAGACCGGGCGACAGTGTGAGCTGGCGGGCTGGCTGACAACTTATCAGAGCC	186
Db	298	CCCCAGAGAGTGCACCCGACACGGTGGGGCTGTACGTCTTTGAGAAAGGATACCATG	357
QY	187	CTGGGGCCCCCTGACTTCCGCAACATGACGGGACTGTGTGACCTTGACACTGTCTCGCAAT	246
Db	358	CTCGACGCGAGGAGCTTTGCGGCGCTGCGGGGCTCGACGCTCTTGAGACTGTGACAGAAC	417
QY	247	GCCATCACCCGATTTGGGGGCCCGGCGCTTTGGGGGACCTCGAGAGCCGTGCCTCCCTCAC	306
Db	418	CAGATGCCAGCTGTGCCAGCGGGGTCTTCCAGCACACTGCGCAACTCAGACAACTTGAGAC	477
QY	307	CTTGAACGGCAACAGGCTGGTGGAGCTGGGCAACGGAGACCTTCGGGGGCCCTGCAATCTG	366
Db	478	CTGACGGCCCAACAGGCTGTGATGAATCAACAATGAGACTTCCGTGGCTCGGGCGCTTC	537
QY	367	CAGCAACCTCATCTCAGCGGGCAACAGAGTGGGCGGATTCGGCGGGAGAGCTTCGAGAC	426
Db	538	GAGCGCTCTTACTTGGGCAAGAACCGCAATCCGCCACATCCAGCTTGATGCTT---GAC	594
QY	427	TTCTTAGAGAGCTTGAAGACCTTGACCTGTCTTACAACAACCTTCGGACAGTTCCTTGG	486
Db	595	ACGCTTGACCGCTCCTGTGAGCTCAAGCTGCAAGGCAACGAGCTGGGGGCACTGCCCG	654
QY	487	GCCGGATGGGGCCATGCTGCTCCTTGACACCTCTAAC	525
Db	655	CTGCGCTGCGCCGCTGCTGTGCTTGAGACTTCAGCCAC	693

```

RESULT 5 -594-6
US-09-706-594-6
: Sequence 6, Application US/09706594
: Patent No. 6670123
: GENERAL INFORMATION:
: APPLICANT: New York Blood Center, Inc.
: APPLICANT: Belyavsky, Alexander
: APPLICANT: Smelkov, Sergey
: APPLICANT: Visser, Jan
: TITLE OF INVENTION: METHOD FOR DETECTING HEMATOPOIETIC STEM CELLS
: FILE REFERENCE: 63475/250
: CURRENT APPLICATION NUMBER: US/09/706,594
: CURRENT FILING DATE: 2000-11-03
: NUMBER OF SEQ. ID NOS: 6
: SOFTWARE: patentIn version 3.1
: SEQ. ID NO 6
: LENGTH: 4250
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-706-594-6

```

Query Match	4.9%;	Score 93.6;	DB 4;	Length 4250;
Best Local Similarity	54.0%;	Pred. No. 4.7e-10;		
Matches 217; Conservative	0;	Mismatches 179;	Indels 6;	Gaps 1;

QY	153	GGAGCTCGGCTGGCTGACAACCTTCAATCCAGGCGCTGGGGCCCCCTGACTTCCGCAACAT	212
Db	1261	GAATAATATCTACGACAGAACTTCATCCGCTGTCTGTGTGGACGACACACTTCTCGAGGC	1320
QY	213	GACGGGACTGTGAGACCTGACACTGTCTGGCAATGCCATACCCGCAATTGGGCGCGCGC	272
Db	1321	CACGGGCGCTGGAACCTCTCGACACTGGGGGAATTAACCGCATCTCGATGATCCAGGACCGCGC	1380
QY	273	CTTTGGGGAACTCGAGAGCTTGGCTTCCCTCCACTTGGACGGCAACAGGCTGTGTGAGACT	332

Db 1381 TTTCGGGGATCTCACCACCTTAAGGCGCTCTAAGTGGCAACAGATGAGAGGCT 1440  
QY 333 GGGACACGGGAGCGCTCCGGGCGCCCGCTCAATCTGCAGACCTCATCTCAGGCAACCA 392  
Db 1441 GAGCCCGGAGTATTATGATGCGCTGCAGAGCTGCAGATCTCTTCTTCACATACATCT 1500  
QY 393 GCTGGGCGCATCGCGCGGAGCGCTTGCAGACATCTTCTAGAGAGCTTGCAGAGCGCTGA 452  
Db 1501 CATCCGCGAGTCTGAGTCTGAACTTTGACCCGGGCCAAGCTCC-----AGCTGCT 1554  
QY 453 CCTGTCCTACACACCTCCGGCAGGTGCGCTGGCGCGCATCGGCGCATCGCTGCT 512  
Db 1555 ATCTGTAATTAACAACCTCTCGAGGCGCATGCGCTTCTGCTTCTGCGCTTACCT 1614  
QY 513 GCACACCTCACTGACCTGACCATACCTTATGACGACCTGCC 554  
Db 1615 CCTCAGGCTAAACCTGAGAGATACCATCTTCACTCTTCTGCC 1656

RESULT 6  
US-08-592-500-1  
; Sequence 1, Application US/08592500  
; Patent No. 6005089  
; GENERAL INFORMATION:  
; APPLICANT: Lanza, Francois  
; APPLICANT: Phillips, David R.  
; APPLICANT: Cazenave, Jean-Pierre  
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/592,500  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/089,455  
; FILING DATE: 09-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Karen B.  
; REGISTRATION NUMBER: 29,684  
; REFERENCE/DOCKET NUMBER: 12418-28  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7452 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: 1462..2419  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2422..4101  
; FEATURE:  
; NAME/KEY: misc signal  
; LOCATION: 68..76

OTHER INFORMATION: //function= "Putitive TPA responsive  
OTHER INFORMATION: element"  
OTHER INFORMATION: /label= TRE  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 471..478  
OTHER INFORMATION: //function= "Ets-1 cis-acting  
OTHER INFORMATION: sequence"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 493..502  
OTHER INFORMATION: //function= "Ets-1 cis-acting  
OTHER INFORMATION: sequence"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: repeat\_region  
LOCATION: 593..881  
OTHER INFORMATION: //rpt\_type= "other"  
OTHER INFORMATION: /label= Alu  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 897..904  
OTHER INFORMATION: //function= "Ets-1 cis-acting  
OTHER INFORMATION: sequence"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: misc binding  
LOCATION: 1142..1149  
OTHER INFORMATION: //function= "Sp1 binding site"  
OTHER INFORMATION: /standard\_name= "Sp1"  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 1178..1184  
OTHER INFORMATION: //function= "Ets-1 cis-acting  
OTHER INFORMATION: sequences"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: TATA\_signal  
LOCATION: 1199..1203  
FEATURE:  
NAME/KEY: TATA\_signal  
LOCATION: 1263..1269  
FEATURE:  
NAME/KEY: misc binding  
LOCATION: 1285..1289  
OTHER INFORMATION: //function= "GATA-1 binding site"  
FEATURE:  
NAME/KEY: misc binding  
LOCATION: 1321..1326  
OTHER INFORMATION: //function= "GATA-1 binding site"  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 1365..1372  
OTHER INFORMATION: //function= "Ets-1 cis-acting  
OTHER INFORMATION: sequences"  
OTHER INFORMATION: /label= Ets-1  
FEATURE:  
NAME/KEY: repeat\_region  
LOCATION: 6133..6440  
OTHER INFORMATION: //rpt\_type= "other"  
OTHER INFORMATION: /label= Alu  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 5610..5615  
OTHER INFORMATION: //standard\_name= "Polyadenylation  
OTHER INFORMATION: signal sequence"  
FEATURE:  
NAME/KEY: misc\_signal  
LOCATION: 6966..6971  
OTHER INFORMATION: //standard\_name= "Polyadenylation  
OTHER INFORMATION: signal sequence"



```

; NAME/KEY: misc_signal
; LOCATION: 7224..7229
; OTHER INFORMATION: /standard_name="Polyadenylation"
; OTHER INFORMATION: signal sequence"
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 7358..7363
; OTHER INFORMATION: /standard_name="Polyadenylation"
; OTHER INFORMATION: signal sequence"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..7452
; OTHER INFORMATION: /standard_name="Nucleotide"
; OTHER INFORMATION: sequence containing the human GPV gene"
US-08-592-500-1

```

```

Query Match 4.8%; Score 91.2; DB 3; Length 7452;
Best Local Similarity 47.8%; Pred. No. 1.6e-09;
Matches 264; Conservative 0; Mismatches 288; Indels 0; Gaps 0;

```

```

QY 273 CTTTGGGACCTTGAGAGCTGGCTTCCCTCCACCTTGAGCGCAAGGCTGGAGCT 332
DB 3276 CTTTGGGAGATGGGGGCGCTTGAGAGCTGGCTGAAACCGACAGCTGGACCTT 3335
QY 333 GGGACACGGGAGGCTCCGGGGCCCGCTCAATCTGACAGCACTCATCTGACGGCAACA 392
DB 3336 GCGCGCGCGCGCTTCCGCACTGAGCGCGCTGCGGTACTTAAAGGGGTACTGTAGCCC 3395
QY 393 GCTGGGCGCGATGCGCGCGGAGCCTTCAAGCACTTCTAAGAGGCTTGAGAGCCTTGA 452
DB 3396 GCGGCTGAGGCGCTTCCGAGGCGCGCTTCCAGGCGCTTGGGAGAGCTGAGCTGCG 3455
QY 453 CCGTGTCTCAACAACACTCCGAGAGTCCCTGGGCGCGATGGGCGCATGCTGCTCCCT 512
DB 3456 CCGTCACTCCACAGGCTGACCGCTCCCGAGCGCTTGTGCGGCGCTCGGCAAGCT 3515
QY 513 GCACACCTCAACCTGAGACCACTTAATGACGACTGCGCCGAGGCGCTTCCGCCA 572
DB 3516 GCGCCAGGATGCTCTGCGCGCGCAAGGCTGCGCGCGCTGCGCGCTTCCGCCA 3575
QY 573 GCTGGTCAAGCTTCCCGCTGAGACTCACTTCAACCGCTGCGCAAGCTGCTCCGA 632
DB 3576 TCTCAGCAGCTGAGAGGCTTCCAGCTCGACCAACAGCTGAGAGCACTGCTGCGCA 3635
QY 633 CCGGCTTTCTCGGAGGCGATGACAGAGCTCTCCCGCGCCCTGGTGTGAGCTT 692
DB 3636 CGTGTGGGCGCTGCTCCCGCTGACGAGAGGCTGTTGGGCGACAACCTCTGCGCGT 3695
QY 693 TAGCGGAAACCCCTGCACTGCAACTGTGAGCTGTGCTGCGGCGGCTGCGCGCG 752
DB 3696 CGACTGTGCGCTGGGCGCTTCTCTGGGATGCTGCGGCGAGCACTTAAAGCTCTG 3755
QY 753 GGAAGAGCTGGAAGAGCTGGGCTTCCCGCGCGCGCTGCGCGCGCTTCTTGGGCA 812
DB 3756 GGAAGAGCGCGCGAGGCTGCGAGCGCTGCGGCGAGCGCGCGCTGCGCTGCGGCG 3815
QY 813 GCGCGAGGCGCA 824
DB 3816 GCGCGAGGCTGA 3827

```

```

RESULT 7
US-08-195-006-1
; Sequence 1, Application US/08195006
; Patent No. 6083688
; GENERAL INFORMATION:
; APPLICANT: Lanza, Francois
; APPLICANT: Phillips, David R.
; APPLICANT: Cazenave, Jean-Pierre
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourile and Crew

```

```

; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/195,006
; FILING DATE: 10-FEB-1994
; CLASSIFICATION: 514
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,455
; FILING DATE: 09-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 12418-28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7452 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1462..2419
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2422..4101
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 68..76
; OTHER INFORMATION: /function="Putative TPA responsive element"
; OTHER INFORMATION: /label="TRE"
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 471..478
; OTHER INFORMATION: /function="Ets-1 cis-acting sequence"
; OTHER INFORMATION: /label="Ets-1"
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 493..502
; OTHER INFORMATION: /function="Ets-1 cis-acting sequence"
; OTHER INFORMATION: /label="Ets-1"
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 593..881
; OTHER INFORMATION: /rpt_type="other"
; OTHER INFORMATION: /label="Alu"
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: 897..904
; OTHER INFORMATION: /function="Ets-1 cis-acting sequence"
; OTHER INFORMATION: /label="Ets-1"
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 1142..1149
; OTHER INFORMATION: /function="Sp1 binding site"
; OTHER INFORMATION: /standard_name="Sp1"

```

```

1 FEATURE:
2 NAME/KEY: misc_signal
3 LOCATION: 1178..1184
4 OTHER INFORMATION: /function= "Ets-1 cis-acting"
5 OTHER INFORMATION: sequences
6 OTHER INFORMATION: /label= Ets-1
7 FEATURE:
8 NAME/KEY: TATA_signal
9 LOCATION: 1199..1203
10 FEATURE:
11 NAME/KEY: TATA_signal
12 LOCATION: 1263..1269
13 FEATURE:
14 NAME/KEY: misc_binding
15 LOCATION: 1285..1289
16 OTHER INFORMATION: /function= "GATA-1 binding site"
17 FEATURE:
18 NAME/KEY: misc_binding
19 LOCATION: 1321..1326
20 OTHER INFORMATION: /function= "GATA-1 binding site"
21 FEATURE:
22 NAME/KEY: misc_signal
23 LOCATION: 1365..1372
24 OTHER INFORMATION: /function= "Ets-1 cis-acting"
25 OTHER INFORMATION: sequences
26 OTHER INFORMATION: /label= Ets-1
27 FEATURE:
28 NAME/KEY: repeat_region
29 LOCATION: 6133..6440
30 OTHER INFORMATION: /rpt_type= "other"
31 OTHER INFORMATION: /label= Alu
32 FEATURE:
33 NAME/KEY: misc_signal
34 LOCATION: 5610..5615
35 OTHER INFORMATION: /standard_name= "Polyadenylation"
36 OTHER INFORMATION: signal sequence"
37 FEATURE:
38 NAME/KEY: misc_signal
39 LOCATION: 6966..6971
40 OTHER INFORMATION: /standard_name= "Polyadenylation"
41 OTHER INFORMATION: signal sequence"
42 FEATURE:
43 NAME/KEY: misc_signal
44 LOCATION: 7224..7229
45 OTHER INFORMATION: /standard_name= "Polyadenylation"
46 OTHER INFORMATION: signal sequence"
47 FEATURE:
48 NAME/KEY: misc_signal
49 LOCATION: 7356..7363
50 OTHER INFORMATION: /standard_name= "Polyadenylation"
51 OTHER INFORMATION: signal sequence"
52 FEATURE:
53 NAME/KEY: misc_feature
54 LOCATION: 1..7452
55 OTHER INFORMATION: /standard_name= "Nucleotide"
56 OTHER INFORMATION: sequence containing the human GPV gene"
57
58 US-08-195-006-1

```

Query Match	4.8%;	Score 91.2;	DB 3;	Length 7452;
Best Local Similarity	47.8%;	Pred. No. 1.6e-09;		
Matches 264;	Conservative 0;	Mismatches 288;	Indels 0;	Gaps 0;

Qy	273	CTTTGGGGGACCTCGAGAGCTTCGCGTTCCTCCGACCTTGAACGGCAACAAGCTCTGGTGGAGCT	332
Db	3216	CTTGGGGGAGATGGGGGGGCTTCGAGAGACTGTGGTGAACCGGACCCGACTGCGACCTT	3335
Qy	333	GGGCAACGGGAGCTTCGCGGGCCCCGTAAATGTGAGACACTCATCTCAGCGGCAACA	392
Db	3336	GCCCCGCGCGCTTCCTCCGCAACTTGAGCGCGCTGTGGTAACTTGAAGGGAGACTCTTGAAGCCC	3395
Qy	393	GCTGGGCGCATCGGCGCGGAGACCTTTCGACGACTTCTTAGAGAGCCTGAGAGACTTGGAA	452
Db	3396	GCGGCTGAGCGCTTTCGCGAGGGGACCTTCAGAGGCTTGTGCGAGCTTCAGAGTGTCTGC	345H

Qy	455	CTGTCTTACAACAACTCTCGGAGAGTCCCTGGGCGCGGATCGGCGCATGCGTGCCT	512
Db	3456	CTGTCACTCCAAACGGCTTGACCGCCCTTCCCGACGGCTTGCTGCGGGCTTCGGCAAGCT	3515
Qy	513	GCACACCTCAACCTGTGACCATTAACCTTATTGACGCACTGCCCCAGGCGCTTGCCCA	572
Db	3516	GGCGCAGGTGCTCTGGCGCGGCAACAGGCTGCGGGCTTGCCCGTGCCTCTTCCGCA	3575
Qy	573	GCTCGTGAAGCTCCCGCTGGACCTCACTCCAAACCGCTGGCCACGCTGCTCCGGA	632
Db	3576	TCTCAGACGCTGGAGAGCGCTCCAGCTCGACCAACACAGCTGGAGACCTCGCTGGGGA	3635
Qy	633	CCGCGTTTCTCTGTGTGGGCTGATGAGAGGCTCTTCCGCGCCCTTGAGTCTGAGCTT	692
Db	3636	CGTTTGTGGGCTTGTGCCCGGCTGAAGGAGGCTCTGTGGGGCAAACTCTGCGCTG	3695
Qy	693	TAGCGGGGACCCCTGCTCACTGCAACTGTGAGCTGCTGTGGCTGGCGGCGCGGC	752
Db	3696	CGACTGTGGCTTGGGGGCTCTTCTGGGGGTGGCTGCGGACAGCACTTAAGCTCTGTGGGCGG	3755
Qy	753	GGACGACCTGGAACGTTGCGCTTCCCGCCCGGCTGAGCGCTACTTCTGAGGAGT	812
Db	3756	GGAAGACGCCCAACGAGTGCAGAGGCTTGGGGCGCAACGCGGCTTGCCGCTCTGAGGCGCT	3815
Qy	813	GCCCGAGGGCGA	824
Db	3816	GCCGAGGGGCTGA	3827

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

/ LOCATION: 6966..6971
/ OTHER INFORMATION: /standard_name= "Polyadenylation"
/ OTHER INFORMATION: signal sequence"
FEATURE:
/ NAME/KEY: misc_signal
/ LOCATION: 7224..7229
/ OTHER INFORMATION: /standard_name= "Polyadenylation"
/ OTHER INFORMATION: signal sequence"
FEATURE:
/ NAME/KEY: misc_signal
/ LOCATION: 7358..7363
/ OTHER INFORMATION: /standard_name= "Polyadenylation"
/ OTHER INFORMATION: signal sequence"
FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1..7452
/ OTHER INFORMATION: /standard_name= "Nucleotide
/ OTHER INFORMATION: sequence containing the human GPV gene"
PCT-US94-07644A-1

Query Match      4.8%; Score 91.2; DB 5; Length 7452;
Best Local Similarity 47.8%; Pred. No. 1.6e-09;
Matches 264; Conservative 0; Mismatches 288; Indels 0; Gaps 0

QY      273      CTTTGGGACCTCGAGACCTCGTTCCTCCATCAGGCAACAGAGCTGTGAAGCT 332
Db      3276      CTTGGGGAGATGGGGGGGCTCGAGAGACTGTGAGTGAACCGCAACGAGCTCGCAACCT 333
QY      333      GGGCACCAGGAGCCTCCGAGGCCCCCGTAATCTGAGACACTCATCTCAAGCGGACCA 392
Db      3336      GCGCGCGCGCCTTCCTCGCAACTGAGCGGCTGGGTAATTAGAGGAGACTTGAAGCC 3395
QY      393      GCTGGAGCGCATCGCGCGGAGAGCCTTTCGACGACTTCTTAGAGAGCTGAGAGACTGGA 452
Db      3396      GGGGTAGAGCGCGCTTCCGAGAGGCGCTTTCAGAGGCTTTCAGAGCTTCAGAGTCTGC 34555
QY      453      CCTGTCTTACAACAACTTCGCGCAGAGTCCCTGGAGCGGCATCGAGCGCATGCTGCCT 512
Db      3456      CCTGACACTCCACAGGCTCGACCGGCTCCCGACGAGCTTGTGCGCGGCTCGGCAGACT 3515
QY      513      GCACACCCCTCAACCGGAGCATTAACCTTATTTAGAGCACTGGCCCCAGGCGGCTTCGCCA 572
Db      3516      GCGCGAGAGTCTCCCTCGCGCGCAAGAGCTCGGCGCTTCGCGCTTCCTTCGCA 35757
QY      573      GCTCGATCACTCTCCGCGCTGACCTCACTCACTCCACACCGCTGGCAAGCTGCTCCGGA 632
Db      3576      TCTCAGCAGCTTGAGAGCGTTCAGCTCGACACACAGCTGAGAGACCTGCTGCGGA 3635
QY      633      CCCGCTTTTCTCTGTGGGCGTGAATGACAGAGGCTCTCCGCGCCCCCTGCTGTGAGCTT 692
Db      3636      CGTGTGGGAGCTCTGCCCCGCGCTGAAGAGAGTCTGTGGGAGGACAACTCTGAGCGCTG 3699
QY      693      TAGCGGAGACCCCTGCACTGCAACTGTGACTGTGCTGTGGCGGCGGCTGGCGGCGC 752
Db      3696      CGACTGTGGGCTTGAGGCGCTTCCTCGAGGAGTGGCTGGAGAGCACTAGAGCTCTGTGGGCG 3755
QY      753      GGAAGACTGGAAGACGTGCGCTCCCGCGCTGGCGCGGCGCTACTTCTGTGGAGT 812
Db      3756      GGAAGACCCCGCAAGTGTGCGAGGACTTGGGCGCAGCGCGGCTGCGCTGTGGGCGCT 3815
QY      813      GCCCGAGGAGCGA 824
Db      3816      GCCGGGGGGTGA 3827

RESULT 9
US-09-799-451-480
: Sequence 480, Application US/09799451
: Patent No. 678369
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Zhou, Ping
: APPLICANT: Goodrich, Ryle

```

APPLICANT: Asundi, Vinod  
APPLICANT: Ren, Feiyan  
APPLICANT: Zhang, Jie  
APPLICANT: Xue, Aidong J.  
APPLICANT: Zhao, Qing A.  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Ma, Yunding  
APPLICANT: Yamazaki, Victoria  
APPLICANT: Chen, Rui-hong  
APPLICANT: Wang, Zhiwei  
APPLICANT: Wang, Dunrui  
APPLICANT: Yang, Yonghong  
APPLICANT: Wehrman, Tom  
APPLICANT: Ghosh, Reena  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. 6783969el Nucleic Acids and  
FILE REFERENCE: Polypeptides  
CURRENT APPLICATION NUMBER: US/09/799,451  
CURRENT FILING DATE: 2001-03-05  
NUMBER OF SEQ ID NOS: 948  
SOFTWARE: pt\_fl\_genes Version 2.0  
SEQ ID NO 480  
LENGTH: 2733  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (271)..(2730)  
US-09-799-451-480

Query Match 4.6%; Score 87.8; DB 4; Length 2733;  
Best Local Similarity 49.7%; Pred. No. 6.4e-09;  
Matches 252; Conservative 0; Mismatches 252; Indels 3; Gaps 1;

QY 289 AGCTGCGTTCCTCCACCTTACCGGCAACAGGCTGCTGAGACTGGGACCGGAGCCTC 348  
DB 436 ACCGAGCAGACCTGGGCTCAACGAGCAAGCTCAAGCCTGCTACTCTTCGCTC 495  
QY 349 CCGGCGCCCGTCAATGTGAGACCTTCAAGGCGCAACCACTGGCCGCGATCGG 408  
DB 496 AACCGCTTTGGAACTCCACCGACTCAACCTTACCAAGACGAGATCTTCTTACATCGAG 555  
QY 409 CCGGAGCCTTGCAGACCTTCTAGAGAGCTGAGAGACTGAGACTGCTGCTTCTTCTCGT 468  
DB 556 GAGGCTGCTTCTCTG---GCCAGTGAAGCTGAGGCTCTGAGCTGGGGCTAACAGAG 612  
QY 469 CTCGCGAGGCTGCTGAGCGGCGATCGGCGCTGCTGCTGCTGCAACCTCAACTG 528  
DB 613 CTCAGCAACCTGACGAGGAGGATGCTGCGAGGATAGCGCTGCAAGTTCTTTTTC 672  
QY 529 GACCAATACCTTATGAGCGACGCGCCGAGGCGCTTGGCCAGCTGGTCAAGCTTCC 588  
DB 673 CAGCAACCTTATGAGGAGTGTGAGCGGCGACCGCTTCTCGAGTGGCCGAGCTCATC 732  
QY 589 CGCTGAGCCTTACCTTCAACCGCTGAGCGAGCTGCTGAGACCGCTTCTTCTCTCT 648  
DB 733 AGATGAGCTGCTTCTTCAACCGCTTCAAGCGCTGAGCGGCTGAGCGGCTTTCGAGCTTC 792  
QY 649 GGGCGTATGAGAGGCGCTTCTCCGCGCTGCTGCTGCTGCTTGAAGCGGAAACCCCTG 708  
DB 793 GCCAGCTGATGCTGTGTGAGCTGCGCGGCAACCCCTTCAACTGTAGTGCAGCTTTC 852  
QY 709 CACTGCACTGTGAGCTGCTGAGCGGCGGCTGAGCGGCGGCGGAGCAACCTGAGAAAG 768  
DB 853 GCGTCTTCTGAGCTGCTGAGGCTTCTTCAACAGCTCAACAGAACTTACGAGCGCTGAG 912  
QY 769 TGGCGCTTCCCGCGCGCGCTGAGCGGCG 795  
DB 913 TGTGAGTCCCGCGGAGGATTTGCCGAGC 939

RESULT 10

US-10-140-002-381  
Sequence 381, Application US/10140002  
Patent No. 6725730  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: ACIDS ENCODING THE SAME  
CURRENT APPLICATION NUMBER: US/10/140,002  
CURRENT FILING DATE: 2002-05-06  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 381  
LENGTH: 2236  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-140-002-381

Query Match 4.6%; Score 87.6; DB 4; Length 2236;  
Best Local Similarity 49.8%; Pred. No. 6.8e-09;  
Matches 222; Conservative 0; Mismatches 224; Indels 0; Gaps 0;

QY 181 CAGGCGCTGGGGCGCCCTGACTTCCGCAACATGACGAGGACTGGTGAACCTGACACTGTCT 240  
DB 649 CAGGCGCTGAGGCTGTGCGCCCTGGGACATCCTGCTGCGAGCGAGCATCTTCTGCAAC 708  
QY 241 CGCAATGCCATCACCCTGATGAGGCGCGGCTTTGGGACCTTGAGAGCTGCGCTTC 300  
DB 709 GGCACCGCATCTCGATGTGCGACCTGCGACCTTCCGCTGCGCAACTCAACATC 768  
QY 301 CTCACCTTGAAGGCAACAGGCTGTGAGAGCTGGGCAACCGGAGCCTCGGGGCGCCCTC 360  
DB 769 CTGTGCTGCACTGCAATGTGCTGTGCGCGAATGTGATGCGGCTTCACTGCGCTGGCC 828  
QY 361 AATGTCAGACCTTCACTTCAAGCGCAACAGCTGAGCGGCGATCGCGCGGAGCTTC 420  
DB 829 CTCCTGAGAGCTGAGCTTCAAGCATATGACAGGCTCCGCTGTGTGAGACCTTGCACA 888  
QY 421 GACGACTTCTTGAAGAGCTGAGGAGCTGAGCTGTCTTCAACAACTTCCGCGAGGTG 480  
DB 889 TTCCAGGCGCTGGGCGGCTTACACAGCTGCACTGAGACCGCTGCGGCTTGAAGAGCTG 948  
QY 481 CCTGGGCGGCGATGAGGCGGCACTGCTGCGCTGCGACACCTTCACTGAGACATACCTT 540  
DB 949 GGCCTGGGCGGCTTCCGCGGCTGAGCTGCTGCGAGTACTTACTTGAAGCAACGCGG 1008  
QY 541 ATTGAGCAGTACCTCCCGAGGCGCTTGGCGAGCTGAGTCACTTCCGCTGAGACTTC 600  
DB 1009 CTGAGGAGCTGCTGATGATACACTTCCGCGAGCTGGGCAACTTACACACTCTTCTC 1068  
QY 601 ACCTCCAACCGCTGAGCAAGCTGAGC 626  
DB 1069 CACGGCAACCGCATCTCCAGCTGCC 1094

RESULT 11  
US-09-461-325-40

```
/ Sequence 40, Application US/09461325A
/ Patent No. 6475753
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: 94 Human Secreted Proteins
/ FILE REFERENCE: P2029P1
/ CURRENT APPLICATION NUMBER: US/09/461,325A
/ PRIOR FILING DATE: 1999-12-14
/ EARLIER APPLICATION NUMBER: 60/089,507
/ EARLIER FILING DATE: 1998-06-16
/ EARLIER APPLICATION NUMBER: 60/089,507
/ EARLIER FILING DATE: 1998-06-16
/ EARLIER APPLICATION NUMBER: 60/089,508
/ EARLIER FILING DATE: 1998-06-16
/ EARLIER APPLICATION NUMBER: 60/089,509
/ EARLIER FILING DATE: 1998-06-16
/ EARLIER APPLICATION NUMBER: 60/089,510
/ EARLIER FILING DATE: 1998-06-16
/ EARLIER APPLICATION NUMBER: 60/090,112
/ EARLIER FILING DATE: 1998-06-22
/ EARLIER APPLICATION NUMBER: 60/090,113
/ EARLIER FILING DATE: 1998-06-22
/ NUMBER OF SEQ ID NOS: 532
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 40
/ LENGTH: 1777
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-461-325-40
```

```
Query Match 4.5%; Score 86; DB 4; Length 1777;
Best Local Similarity 49.6%; Pred. No. 1,4e-08;
Matches 221; Conservative 0; Mismatches 225; Indels 0; Gaps 0;
```

```
QY 181 CAGGCGCTGGGGCCCCCTGACTTCGCAACATGACGGAGACTGGTGAACCTGACATGTCT 240
DB 174 CAGGCGCTGAGGCTGTGCCCCGGGAGCATCTGCTGCCAGCCAGCGCATCTTCTGAC 233
QY 241 CGCAATGCCATCACCGCATTTGGGGCCCGGCTTTGGGGGACCTCGAGAGCTGCGTTC 300
DB 234 GCGAACCGCATCTCGATGTGCCAGCTGCCAGCTTCCGTGCTGCCGCACTTCCACCATC 293
QY 301 CTCACCTTGACGGCAACAGGCTGTGAGAGCTGGGACCGGAGACCTCCGGGGCCCCGTC 360
DB 294 CTGTGGCTGACCTCGAATGTGCTGGCCGCAATTGATGGCGCTGCTTCACTGGCTGGGC 353
QY 361 AATCTGACAGCACTCTATCTTCAAGCGGCAACAGCTGGGCTGCATCGCGCGGAGCTTC 420
DB 354 CTCCTGAGAGAGTGAAGCTTCAAGCAAGCTCGGATCTGTGTGACCTTGCACCA 413
QY 421 GAGCACTTCTAGAGAGCTGAGAGCACTGACCTGTGCTTCAACAACCTCCGGGAGGTG 480
DB 414 TTCACGCGCTGGGCGGCTTACACAGGTGCACTTGAACCGCTGCGGCTTGCAGAGCTG 473
QY 481 CCTTGGGCGGCGATCGGGCGCATGCTGCTGCTGACACCTTCAACCTGAGCACTTAACCTT 540
DB 474 GGGCCGCGGCTGTTCGCGGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 533
QY 541 ATTGAGCACTGCGCCCGAGCGCTTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 534 CTGACGAGCACTGCTGATGACATCTTCCGAGACTGGGCAACCTTCAACACCTTCTTCTG 593
QY 601 ACCTCCACCGCTGCGCAAGCTGGC 626
DB 594 CACGGCAACCGCATCTCCAGCGTGC 619
```

```
RESULT 12
US-10-012-542-40
/ Sequence 40, Application US/10012542
/ Patent No. 6627741
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
```

```
/ TITLE OF INVENTION: 94 Human Secreted Proteins
/ FILE REFERENCE: P2029P1
/ CURRENT APPLICATION NUMBER: US/10/012,542
/ PRIOR FILING DATE: 2001-12-12
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
/ PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
/ PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
/ PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
/ PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
/ PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
/ PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
/ PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
/ NUMBER OF SEQ ID NOS: 532
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 40
/ LENGTH: 1777
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-012-542-40
```

```
Query Match 4.5%; Score 86; DB 4; Length 1777;
Best Local Similarity 49.6%; Pred. No. 1,4e-08;
Matches 221; Conservative 0; Mismatches 225; Indels 0; Gaps 0;
```

```
QY 181 CAGGCGCTGGGGCCCCCTGACTTCGCAACATGACGGAGACTGGTGAACCTGACATGTCT 240
DB 174 CAGGCGCTGAGGCTGTGCCCCGGGAGCATCTGCTGCCAGCCAGCGCATCTTCTGAC 233
QY 241 CGCAATGCCATCACCGCATTTGGGGCCCGGCTTTGGGGGACCTCGAGAGCTGCGTTC 300
DB 234 GCGAACCGCATCTCGATGTGCCAGCTGCCAGCTTCCGTGCTGCCGCACTTCCACCATC 293
QY 301 CTCACCTTGACGGCAACAGGCTGTGAGAGCTGGGACCGGAGACCTCCGGGGCCCCGTC 360
DB 294 CTGTGGCTGACCTCGAATGTGCTGGCCGCAATTGATGGCGCTGCTTCACTGGCTGGGC 353
QY 361 AATCTGACAGCACTCTATCTTCAAGCGGCAACAGCTGGGCTGCATCGCGCGGAGCTTC 420
DB 354 CTCCTGAGAGAGTGAAGCTTCAAGCAAGCTCGGATCTGTGTGACCTTGCACCA 413
QY 421 GAGCACTTCTAGAGAGCTGAGAGCACTGACCTGTGCTTCAACAACCTCCGGGAGGTG 480
DB 414 TTCACGCGCTGGGCGGCTTACACAGGTGCACTTGAACCGCTGCGGCTTGCAGAGCTG 473
QY 481 CCTTGGGCGGCGATCGGGCGCATGCTGCTGCTGACACCTTCAACCTTGAACCTTAACCTT 540
DB 474 GGGCCGCGGCTGTTCGCGGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 533
QY 541 ATTGAGCACTGCGCCCGAGCGCTTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 534 CTGACGAGCACTGCTGATGACATCTTCCGAGACTGGGCAACCTTCAACACCTTCTTCTG 593
QY 601 ACCTCCACCGCTGCGCAAGCTGGC 626
DB 594 CACGGCAACCGCATCTCCAGCGTGC 619
```

```
RESULT 13
US-10-115-123-40
/ Sequence 40, Application US/10115123
/ Patent No. 6774216
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: 94 Human Secreted Proteins
/ FILE REFERENCE: P202930A1D2
/ CURRENT APPLICATION NUMBER: US/10/115,123
/ CURRENT FILING DATE: 2002-04-04
```

```

; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 1777
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-115-123-40

```

```

Query Match      4.5%; Score 86; DB 4; Length 1777;
Best Local Similarity 49.6%; Pred. No. 1.4e-08;
Matches 221; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

```

```

QY 181 CAGGCGCTGGGCCCCCTGACTTCGCAATGACGGGACTGTGAGACCTGACACTGTCT 240
DB 174 CAGGCGCTGGGCCCCCTGACTTCGCAATGACGGGACTGTGAGACCTGACACTGTCT 233
QY 241 CGCAATGCCATCACCCTGATTTGGGGCCGCGCTTTGGGGACCTGAGAGCCTGGCTTC 300
DB 234 GGCACCCGATCTCGCATGTGCGAGCTGCGACCTTCGCTGCGCGCAACTCAACATC 293
QY 301 CTCACCTTGAAGGCAACAGGCTGTGAGCTGGGACCGGAGCCTCCGGGGCCCGCTC 360
DB 294 CTGTGGCTGCACTCGAATGTGTGCGCCGCAATTGATGCGGCTTCACTGGCCCTGGCC 353
QY 361 AATTCGACGACCTTCACTCTCAGCGGCAACGAGCTGGGCGGATGCGCGCGGAGGCTTC 420
DB 354 CTCCTGAGAGACTGAGACTTCAGCGGATGACAGCTCTGCTGTGTGAGACCTTCGAC 413
QY 421 GACGACTTCTTGAAGAGCTTGAAGAGCTGAGACTGTCTCTTCAACAACTTCGGGAGTG 480
DB 414 TTCCAGCGGCTGGGCGGCTTACACAGGTGACCTGAGACCGCTGGGCGCTGAGGAGCTG 473
QY 481 CCTGGGCGGCGATGAGGCGGCTGCTGCTGCTGACACCTTCAACTTGAACATACTT 540
DB 474 GGCCTGGGCTGTTCGCGGCTGTGCTGCTGCTGCTGCTTACTTGAAGAGCAACGCG 533
QY 541 ATTGAGCACTGCGGCGGCTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 534 CTGCAAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 593
QY 601 ACCTTCACCGGCTGGGCGGCTGGC 626
DB 594 CACGGCAACCGCATCTCCAGCGTGCC 619

```

```

RESULT 14
US-09-520-781-9
; Sequence 9, Application US/09520781
; Patent No. 6689866
; GENERAL INFORMATION:
; APPLICANT: Shimketo, Richard A.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
; FILE REFERENCE: 15966-540 No. 6689866el Polynucleotides
; CURRENT APPLICATION NUMBER: US/09/520,781
; CURRENT FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: USSN 60/123,667
; PRIOR FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0

```

```

; SEQ ID NO 9
; LENGTH: 2341
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (215)..(2173)
US-09-520-781-9

```

```

Query Match      4.2%; Score 79.6; DB 4; Length 2341;
Best Local Similarity 51.1%; Pred. No. 2.8e-07;
Matches 187; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

```

```

QY 58 CCTGCGTGTGCGCAACCTTGTCCGAGTGTGCTGACAGCCCTTGTGGCCACCGAGGCTG 117
DB 353 CCTGCGTGTGCTGTGTGAGTAACCACTTCAAGAGGTGTGTGACCGCGGCGCTTC 412
QY 118 CTGTTTGTGCGCCCAAGTGTGAGCCGCGGCAAGTGTGAGCTGGCTGTGACAACTTC 177
DB 413 TCCGAGTGTCCGCAAGGTATTCCTCGAACACCGGTAACCTCACTATGAGAAACAC 472
QY 178 ATCCAGGCGCTGGGCCCCCTGACTTCGCAACATGAGGAGCTGTGACTGTGACTG 237
DB 473 ATCCAGATGATTCAGGCGGACACCTTCGCACTTCAACACCTGAGGTGTGAGTGTG 532
QY 238 TCTGCAATGACATCACCCGATTTGGGCGCGGCTTTGGGAGACCTGAGAGCCTGCGT 297
DB 533 GGCAGGAATCTCATCCGCGAGATTGAGTGGGCGCTTCACAGGCTGGCGAGCTCAGC 592
QY 298 TCCCTCACTTGAAGGCAACAGGCTGTGAGCTGGGACCGGAGGCTCCGGGGCCCG 357
DB 593 ACCCTGAGAGCTGTGCAACATGTGCTGACATCTCCAGCGGCGCTTTGATTAATCTG 652
QY 358 GTCAATGTGAGCACTTCACTCTCAGCGGCAACGAGTGGGCGGATGCGCGCGGAGGC 417
DB 653 TCCAGCTGCGGAGCTGTGCTGTGCAACACCCATGAAAGATCTCTTACGCGC 712
QY 418 TTCGAC 423
DB 713 TTCAAC 718

```

```

RESULT 15
US-09-520-781-11
; Sequence 11, Application US/09520781
; Patent No. 6689866
; GENERAL INFORMATION:
; APPLICANT: Shimketo, Richard A.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
; FILE REFERENCE: 15966-540 No. 6689866el Polynucleotides
; CURRENT APPLICATION NUMBER: US/09/520,781
; CURRENT FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: USSN 60/123,667
; PRIOR FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2607
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (215)..(1984)
US-09-520-781-11

```

```

Query Match      4.2%; Score 79.6; DB 4; Length 2607;
Best Local Similarity 51.1%; Pred. No. 2.8e-07;
Matches 187; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

```

```

QY 58 CCTGCGTGTGCGCAACCTTGTCCGAGTGTGCTGACAGCCCTTGTGGCCACCGAGGCTG 117
DB 353 CCTGCGTGTGCTGTGTGAGTAACCACTTCAAGAGGTGTGTGACCGCGGCGCTTC 412

```

```
QY 118 CTGTTGTGCGGCCCAAGTGAACCGGCGACAGTGAAGCTGGGCTGAGCAACTTC 177
Db 413 TCCGAGGTCCCGCAGGGTATTCCTTGACACCCGGTACTCAACCTCATGGAGAACAC 472
QY 178 ATCCAGGCCCTGGGGCCCTCTGAATTCCGACATGACGGACTGTGACTGACACTG 237
Db 473 ATCCAGATGATCCAGGCCGACACCTTCGGCACCTCCACCACTGGAGGTCTGCAATTG 532
QY 238 TCTCGCAATGCCATCACCAGCATTTGGGGCCCGGCTTTGGGACCTCGAAGCTGGCT 297
Db 533 GGCAGGAACCTCCATCCGGCAGATTGAAGTGGGGCCCTTCACGCGCTGGCCAGCTCAGC 592
QY 298 TCCCTCCACCTTGACGGGCAACAGGCTGATGAGCTGGGCACCGGAGACCTCCGGGGCCCC 357
Db 593 ACCCTGGAGCTGTTCGACAACTGGCTGACAGTCATCCCTTAGCGGGCTTTGAAATACCTG 652
QY 358 GTCAATCTGAGCACTCATCTCAGGGGCAACCAAGCTGGGCCGCAATCGGCCGGGAGCC 417
Db 653 TCCAGCTGGGGAGCTGTGGCTTCGCAACAAACCCCATGAAAGCATCCCTCTTAAGCC 712
QY 418 TTGAC 423
Db 713 TTCAAC 718
```

Search completed: November 20, 2004, 14:51:20  
Job time : 219 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 06:40:05 ; Search time 12916 Seconds  
(without alignments)  
6985.812 Million cell updates/sec

Title: US-10-071-879-9  
Perfect score: 1908  
Sequence: 1 atgccccgcgcctcctcgtc.....tgaagagagatgtgtgtga 1908

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1908	100.0	2449	9	BC015581 Homo sapi
2	1665	87.3	1665	6	CO718542 Sequence
3	1629	85.4	1653	6	AX358336
4	1585.4	83.1	1989	9	BC027475
5	1525.4	79.9	2553	10	BC023036
6	1523.8	79.9	2612	10	BC023156
7	1355	71.0	139072	9	AP000485
8	1355	69.4	180135	2	AC018798 Homo sapi
9	1335	68.9	180135	2	AC018798 Homo sapi
10	1314	68.9	180135	2	AC018798 Homo sapi
11	1185	60.5	1571	9	BC071866
12	1108.2	58.1	238317	2	AC119332
13	1101.8	57.7	1653	6	AX358338
14	1083.8	56.8	269255	2	AC106165
15	1082.2	56.7	195342	10	AC128739
16	992	52.0	1288	9	BC014040
17	985	51.6	1286	6	BC000207
18	624.8	32.7	1887	6	AX179258
19	624.8	32.7	1887	6	AX266997

20	624.8	32.7	2067	9	AK172754	AK172754 Homo sapi
21	624.8	32.7	2549	6	BC003578	BC003578 Homo sapi
22	624.8	32.7	2558	6	AX266995	AX266995 Sequence
23	621	32.5	3144	9	AB033072	AB033072 Homo sapi
24	619.4	32.5	2818	6	AR338601	AR338601 Sequence
25	617.4	32.4	3320	6	AB056799	AB056799 Macaca fa
26	617	32.3	2316	6	CQ730862	CQ730862 Sequence
27	617	32.3	3386	6	AX675281	AX675281 Sequence
28	617	32.3	3568	6	AX675285	AX675285 Sequence
29	615.4	32.3	1734	6	CO721191	CO721191 Sequence
30	612	32.1	2855	6	AX574572	AX574572 Sequence
31	612	32.1	2855	6	AT358127	AT358127 Homo sapi
32	591.2	31.0	2450	6	AX675283	AX675283 Sequence
33	583.4	30.6	2911	10	AK122476	AK122476 Mus muscu
34	552.4	29.0	3174	6	AX676904	AX676904 Sequence
35	552.4	29.0	3174	9	AB040917	AB040917 Homo sapi
36	549.2	28.8	2831	10	BC066999	BC066999 Mus muscu
37	518.8	27.2	138538	9	AC011445	AC011445 Homo sapi
38	518.8	27.2	172090	2	AC018477	AC018477 Homo sapi
39	518.6	27.2	92797	9	AF038458	AF038458 Homo sapi
40	511	26.8	100426	9	AL591063	AL591063 Human DNA
41	488.2	25.6	159443	2	AC101792	AC101792 Mus muscu
42	488.2	25.6	221980	10	AC124583	AC124583 Mus muscu
43	487	25.5	247615	2	AC098061	AC098061 Rattus no
44	487	25.5	277055	2	AC127880	AC127880 Rattus no
45	482.4	25.3	179820	2	AC011679	AC011679 Homo sapi

## ALIGNMENTS

RESULT 1  
LOCUS BC015581  
DEFINITION Homo sapiens leucine rich repeat and fibronectin type III domain containing 4, mRNA (cDNA clone MGC:23278 IMAGE:4637053), complete cds.  
ACCESSION BC015581  
VERSION BC015581.2 GI:33874560  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Wax, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullen, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hilly, S.W., Villalón, D.X., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, J., Small, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

TITLE  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
PUBMED 12477932  
REFERENCE Strausberg, R.  
AUTHORS Direct Submission  
TITLE Submitted (04-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590.







Db 961 TTGTTGGCAACTCTCCCGAGCCGGGCTTTCCCAACGGGACCTTAGAGATTGGGGTG 1020  
 QY 1021 ACCGGGCGCTGGGGAGCGCTGGGGGCTAAACCTGCATTCGCCAACCACTTGTGTAAGCC 1080  
 Db 1021 ACCGGGCGCTGGGGAGCGCTGGGGGCTAAACCTGCATTCGCCAACCACTTGTGTAAGCC 1080  
 QY 1081 ACAGCCCGATAGAACTGCGGGTGCTGGCTTTGCCCATGGTGGGAACAGAGTGGCCAG 1140  
 Db 1081 ACAGCCCGATAGAACTGCGGGTGCTGGCTTTGCCCATGGTGGGAACAGAGTGGCCAG 1140  
 QY 1141 GGGGGCGCGCCCGCGGCTCGAACATCGCGCTCCCTCGCACTGCTGCCAGAGGTGAG 1200  
 Db 1141 GGGGGCGCGCCCGCGGCTCGAACATCGCGCTCCCTCGCACTGCTGCCAGAGGTGAG 1200  
 QY 1201 GGGAGCTGGAGTGTGAGCCAGCGGTGACAGGTGACCGGACCTTCCAAATCCAGTACAC 1260  
 Db 1201 GGGAGCTGGAGTGTGAGCCAGCGGTGACAGGTGACCGGACCTTCCAAATCCAGTACAC 1260  
 QY 1261 GTGAGCTGGAGTGTGAGCCAGCGGTGACAGGTGACCGGACCTTCCAAATCCAGTACAC 1320  
 Db 1261 GTGAGCTGGAGTGTGAGCCAGCGGTGACAGGTGACCGGACCTTCCAAATCCAGTACAC 1320  
 QY 1321 AGCAGGGAATAGAACCTTCACTTACCGGATTGTCCAGCTTCCAGCCACACTTCTTG 1380  
 Db 1321 AGCAGGGAATAGAACCTTCACTTACCGGATTGTCCAGCTTCCAGCCACACTTCTTG 1380  
 QY 1381 CTGAAGCACTCGTCCCGCGGCGTGAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
 Db 1381 CTGAAGCACTCGTCCCGCGGCGTGAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
 QY 1441 GCTGGGCGCTCTGACCTCAAGGACCAAGGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
 Db 1441 GCTGGGCGCTCTGACCTCAAGGACCAAGGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
 QY 1501 CGGCGCTCGCTCTGAGCCACGCTTCAAGGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
 Db 1501 CGGCGCTCGCTCTGAGCCACGCTTCAAGGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
 QY 1561 GCGGTGGGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
 Db 1561 GCGGTGGGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
 QY 1621 GGGCGGGGGGGCGGAATGCGCGCTCCCTCAAGCTCAGCCAC 1665  
 Db 1621 GGGCGGGGGGGCGGAATGCGCGCTCCCTCAAGCTCAGCCAC 1665

RESULT 3  
 AX358336 1653 bp DNA linear PAT 13-FEB-2002

LOCUS AX358336 1653 bp DNA linear PAT 13-FEB-2002  
 DEFINITION Sequence 7 from Patent WO0202604.  
 ACCESSION AX358336  
 VERSION AX358336.1 GI:18674972  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 Thayer E.C., Sheppard P.O. and Presnell S.R.  
 Leucine-rich repeat proteins, zlr7, zlr8 and zlr9  
 Patent: WO 0202604-A 7 10-JAN-2002;  
 JOURNAL ZymoGenetics, Inc. (US)  
 FEATURES  
 source 1..1653  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 1..>1653  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD23323.1"  
 /db\_xref="GI:18674973"

/translation="MAPPLLLILASGAACPILVCQNLSESLTCAHRLGLFVPP  
 NVRRIVELALANFIQALGPDPRMTGTVDLTLLRNALITRIGARAFGLBLSRSH  
 LDNRIVLEGTGSLRGVNLQHLILSGNQLGRJAPGAFDPLSLDLDSYNNRKY  
 PMAGIAMPALHTLNLIDALPGAFQGLSRLDLTSRLTLADPLPESGR  
 DAPASPAVLVSPSGNDHCHCNLWLRRLARPDDLETCSPPGLANRYWAPDEPF  
 SCEPPLIARHTOBLMYLEGORATLRALDPAPTMHWVSPDRIVNRRARAFNG  
 TLEIGTGGADDAAGTYTCTATNPAGENTARVELLVALLPHGNSASABGRPSDILAS  
 ARTAASBEGTLESEPAVQVEVATSGLVSWGGRPADPVMFOIQVNSSEDTLYR  
 IVPASSHHFLKHLVPGADYDLCLILASPAAGSDILATRLGCAHSTLPASPLCHA  
 LQAHVLGTLTVAVGVVLAITLHCGPCWFGAGPEMAASP"

## ORIGIN

Query Match 85.4%; Score 1629; DB 6; Length 1653;  
 Best Local Similarity 99.9%; Pred. No. 1,1e-216;  
 Matches 1651; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 ATGGCCCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60  
 Db 1 ATGGCCCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60  
 QY 61 TGGCTTGCAGAACTGTTCGAGTGTGCTGAGCACTCTGTGCCACCGAGGCTGTG 120  
 Db 61 TGGCTTGCAGAACTGTTCGAGTGTGCTGAGCACTCTGTGCCACCGAGGCTGTG 120  
 QY 121 TTGTGCGCGCCCAACGTTGACCCGCGCACAGTGAAGCTGGCTGAGCACTTCAATC 180  
 Db 121 TTGTGCGCGCCCAACGTTGACCCGCGCACAGTGAAGCTGGCTGAGCACTTCAATC 180  
 QY 181 CAGGCGCTGGGGCGCCCTGACTTCCGCAACATGACGGGACTGGTGAACCTGACACTGTCT 240  
 Db 181 CAGGCGCTGGGGCGCCCTGACTTCCGCAACATGACGGGACTGGTGAACCTGACACTGTCT 240  
 QY 241 CGCAATGCCATCAACCGCATTTGGGAGCGCGCTTTGGGAGACTGAGAGCCCTGCTTC 300  
 Db 241 CGCAATGCCATCAACCGCATTTGGGAGCGCGCTTTGGGAGACTGAGAGCCCTGCTTC 300  
 QY 301 CTCACCTTGAAGGCAACAGGCTGTGAGCTGGGAGACCGGAGACCTCCGGGGCCCCGTC 360  
 Db 301 CTCACCTTGAAGGCAACAGGCTGTGAGCTGGGAGACCGGAGACCTCCGGGGCCCCGTC 360  
 QY 361 AATCTGACACACCTCATCTTCAGCGGCAACAGCTGGGCGCATCGCGCGGGAGCCTTC 420  
 Db 361 AATCTGACACACCTCATCTTCAGCGGCAACAGCTGGGCGCATCGCGCGGGAGCCTTC 420  
 QY 421 GACGACTTCTTGAAGAGCTTGAAGACCTTGAACCTTCTTCAACAACACTCCGGCAGGTG 480  
 Db 421 GACGACTTCTTGAAGAGCTTGAAGACCTTGAACCTTCTTCAACAACACTCCGGCAGGTG 480  
 QY 481 CCTGGGCGGGCATGCGGCGCATGCTGCTGCGACACCTGCAACCTGGAACCATACCTT 540  
 Db 481 CCTGGGCGGGCATGCGGCGCATGCTGCTGCGACACCTGCAACCTGGAACCATACCTT 540  
 QY 541 ATTGAAGCACTGCCCCCAAGCGCTTTCGCGCAAGCTGCTCACTCTCCGCTGAGACTC 600  
 Db 541 ATTGAAGCACTGCCCCCAAGCGCTTTCGCGCAAGCTGCTCACTCTCCGCTGAGACTC 600  
 QY 601 ACTTCAACCGCTTGGCCACAGCTGCTCGAACCCGCTTTCTCTGCGGGCGTGAATGA 660  
 Db 601 ACTTCAACCGCTTGGCCACAGCTGCTCGAACCCGCTTTCTCTGCGGGCGTGAATGA 660  
 QY 661 GAGGCTCTCCCGCGCCCTGCTGAGCTTGAAGGGGAACCCCTGACAGCAACTGT 720  
 Db 661 GAGGCTCTCCCGCGCCCTGCTGAGCTTGAAGGGGAACCCCTGACAGCAACTGT 720  
 QY 721 GAGTGTGTGCTGCGGCGGCTGAGCGCGCGGACGACTTGAAGAACTGCGCTCCCGC 780  
 Db 721 GAGTGTGTGCTGCGGCGGCTGAGCGCGCGGACGACTTGAAGAACTGCGCTCCCGC 780  
 QY 781 CCGGCGCTGGCGCGCGCTTACTTCTTGGGCAATGCGCGGCGAGTTCTCTGTGAGCGG 840  
 Db 781 CCGGCGCTGGCGCGCGCTTACTTCTTGGGCAATGCGCGGCGAGTTCTCTGTGAGCGG 840  
 QY 841 CCGCTCATTTGCCCGGCAACGAGAGGCTTGGGGTGTGGAAGGCAACCGGGCCACGCTG 900

```

Db      |CCCCCATATGCGCCGACACGACGAGCGCTCTGGGTGCTGGAAAGCCAGCGGCGCACGCTG 900
Qy      |901 CCGTCCCGGCGCCCTGGGTGACCCCGGCTTACCAATGCACTGGGTGGTCTTGACACACCGG 960
Db      |901 CCGTCCCGGCGCCCTGGGTGACCCCGGCTTACCAATGCACTGGGTGGTCTTGACACACCGG 960
Qy      |961 TTGGTGGGCACTCCCTCCGAGCCCGGGCTTCCCAACGAGGACCTTGAAGATGGGGTG 1020
Db      |961 TTGGTGGGCACTCCCTCCGAGCCCGGGCTTCCCAACGAGGACCTTGAAGATGGGGTG 1020
Qy      |1021 ACCGCGCTGGGAGCGCTGGGGGCTACACCTTCATGCGCACCAACCTCTGGTGAAGGCC 1080
Db      |1021 ACCGCGCTGGGAGCGCTGGGGGCTACACCTTCATGCGCACCAACCTCTGGTGAAGGCC 1080
Qy      |1081 ACAGCCCGAGTGAAGATCGCGGGGTGCTGGCTTGGCCCATGGTGGGAAACAGCAGTCCGAG 1140
Db      |1081 ACAGCCCGAGTGAAGATCGCGGGGTGCTGGCTTGGCCCATGGTGGGAAACAGCAGTCCGAG 1140
Qy      |1141 GGGGGGCGGCGCGGCGGCTCGGACATCGCGCGCTCGCTCGCACTGGCTGCGAGGGGTAG 1200
Db      |1141 GGGGGGCGGCGCGGCGGCTCGGACATCGCGCGCTCGCTCGCACTGGCTGCGAGGGGTAG 1200
Qy      |1201 GGGAGCGCTGAGTCTGAGCGACCGCGTGAAGTGAAGAGTGAAGCGGACCTTCAAGGGCTG 1260
Db      |1201 GGGAGCGCTGAGTCTGAGCGACCGCGTGAAGTGAAGAGTGAAGCGGACCTTCAAGGGCTG 1260
Qy      |1261 GTGAGCTGGGGGTCCCGGGGCGGACCGCGTGAAGTGAAGAGTGAAGCGGACCTTCAAGGGCTG 1320
Db      |1261 GTGAGCTGGGGGTCCCGGGGCGGACCGCGTGAAGTGAAGAGTGAAGCGGACCTTCAAGGGCTG 1320
Qy      |1321 AGCAGGAGATGAGACCTCTCATCTACCGGATTTGCTCCAGCTCCAGCACCACCTTCCG 1380
Db      |1321 AGCAGGAGATGAGACCTCTCATCTACCGGATTTGCTCCAGCTCCAGCACCACCTTCCG 1380
Qy      |1381 CTGAAGCACCTGCTCCCGCGCGCTGACATGACCTTGTGCTGCTGCTGCTGCTGCTGCTG 1440
Db      |1381 CTGAAGCACCTGCTCCCGCGCGCTGACATGACCTTGTGCTGCTGCTGCTGCTGCTGCTG 1440
Qy      |1441 GCTGGGCGCTCTGACCTCAACGCGCACAGGCTGCTGGGCTGCTGGCCATTTCTCCAGCTG 1500
Db      |1441 GCTGGGCGCTCTGACCTCAACGCGCACAGGCTGCTGGGCTGCTGGCCATTTCTCCAGCTG 1500
Qy      |1501 CCGGCGCTCGCGCTGTCACCGCGCTGACAGGCGCCACGCTGCTGGGCGGAGCCCTGACCGTG 1560
Db      |1501 CCGGCGCTCGCGCTGTCACCGCGCTGACAGGCGCCACGCTGCTGGGCGGAGCCCTGACCGTG 1560
Qy      |1561 GCGGTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1618
Db      |1561 GCGGTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1618
Qy      |1619 GGGGCGGGGGGCGGGAATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1651
Db      |1619 GGGGCGGGGGGCGGGAATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1651

```

```

RESULT 4
BC027475      1889 bp      mRNA      linear      PRI 04-JUN-2004
LOCUS      BC027475
DEFINITION      Homo sapiens leucine rich repeat and fibronectin type III domain
ACCESSION      BC027475
VERSION      BC027475.1
KEYWORDS      GI:22268152
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens

```

```

REFERENCE
AUTHORS      Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
              Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schliefer, G.D.,
              Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
              Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

```

```

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loguélano, N.A., Peters, G.J.,
Abramson, R.D., Mullen, S.D., Bosk, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huliy, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Hellon, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Boulford, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1889)
Strausberg, R.
Direct Submission
Submitted (04-Apr-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgapsb-remail.nih.gov
Tissue Procurement: DCTD/DRP/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Huliy, S.W., Loulseg, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

```

## FEATURES

source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
 Series: IRAX Plate: 52 Row: a Column: 22.

## gene

1..1889

Location/Qualifiers

1..1889

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3887234"

/tissue\_type="lung, carcinoma, large cell undifferentiated."

/clone\_1lb="NIH\_MGC\_69"

/lab\_host="DH10B"

/note="vector: PCMV-SPORT6"

<1..1889

/gene="LRFN4"

/note="synonym: MGC3103"

/db\_xref="LocusID:78999"

<1..1587

/gene="LRFN4"

/codon\_start=1

/product="LRFN4 protein"

/protein\_id="AAH27475.2"

/db\_xref="GI:48257123"

/db\_xref="LocusID:78999"

/translation="LVECTGSLRGVNLQHLILSGNOLGRIAPGAFDDFLESLIEDLISYNIRQVPAGIGAMPALHTINDPHNLIDLPAPAFQOLGSLDLSNLTATLAPDLSRGRDADASPVLVSSGNFVLCNCELMRLRLRPDDLTCSPPCLAGRYFMAVEGEFSCPEPLIARTQRLWLEGGRAFLRCALDPAFTMHWVGPDDLVGSSARAFPNGLTLEIGATGADAGATCTIATNPAGAVARVELRTALPHGSSAEGGRSPSDIATAGFAARAEGETLSEPAVQTEVATSGLVSGRPAADPYMMFOIYNSSDRLTIYIVASSHPLFLKLIVPADVDICLIALSPAGSPDLTATRLICAHFSTLPASPLCALQAHVIGLITTVAVGVVALLVFTVALLVRGAGNGRLPLKLSHYQ

## CDS

1..1889

Location/Qualifiers

1..1889

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3887234"

/tissue\_type="lung, carcinoma, large cell undifferentiated."

/clone\_1lb="NIH\_MGC\_69"

/lab\_host="DH10B"

/note="vector: PCMV-SPORT6"

<1..1889

/gene="LRFN4"

/note="synonym: MGC3103"

/db\_xref="LocusID:78999"

<1..1587

/gene="LRFN4"

/codon\_start=1

/product="LRFN4 protein"

/protein\_id="AAH27475.2"

/db\_xref="GI:48257123"

/db\_xref="LocusID:78999"

/translation="LVECTGSLRGVNLQHLILSGNOLGRIAPGAFDDFLESLIEDLISYNIRQVPAGIGAMPALHTINDPHNLIDLPAPAFQOLGSLDLSNLTATLAPDLSRGRDADASPVLVSSGNFVLCNCELMRLRLRPDDLTCSPPCLAGRYFMAVEGEFSCPEPLIARTQRLWLEGGRAFLRCALDPAFTMHWVGPDDLVGSSARAFPNGLTLEIGATGADAGATCTIATNPAGAVARVELRTALPHGSSAEGGRSPSDIATAGFAARAEGETLSEPAVQTEVATSGLVSGRPAADPYMMFOIYNSSDRLTIYIVASSHPLFLKLIVPADVDICLIALSPAGSPDLTATRLICAHFSTLPASPLCALQAHVIGLITTVAVGVVALLVFTVALLVRGAGNGRLPLKLSHYQ

SOJNGSPPTKAPHPSPPPRQSCSLDLSGACGYARLUGAMARSHVHGL  
LGAGCRGVGGSABRLEBSVY"

Query Match 83.1%; Score 1585.4; DB 9; Length 1889;  
Best Local Similarity 99.9%; Pred. No. 1.1e-210;  
Matches 1586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 322 CTGTGGAGCTGGGACCGGAGGCTCCGGGGGCCCCGTCATCTGACGACCTCATCTC 381  
Db 1 CTGTGGAGCTGGGACCGGAGGCTCCGGGGGCCCCGTCATCTGACGACCTCATCTC 60  
QY 382 AGCGGACACCAAGCTGGGCGCCATGCGCGCGGAGCTTCGACGACTTCCTAGAGGCTG 441  
Db 61 AGCGGACACCAAGCTGGGCGCCATGCGCGCGGAGCTTCGACGACTTCCTAGAGGCTG 120  
QY 442 GAGGACCTGGAAGCTGCTGCTAACAACACTTCGCGAGGTGCTGGGCGGAGCTGAGGCGC 501  
Db 121 GAGGACCTGGAAGCTGCTGCTAACAACACTTCGCGAGGTGCTGGGCGGAGCTGAGGCGC 180  
QY 502 ATGCGTGGCCCTGACACCTCAACCTGGAACATTAACCTTAATGACGACTGCGCCGACG 561  
Db 181 ATGCGTGGCCCTGACACCTCAACCTGGAACATTAACCTTAATGACGACTGCGCCGACG 240  
QY 562 GCCTTGGCCAGCTCGGTGAGCTCTCCGCTGAGACTCACTCAACCGCTGAGCAG 621  
Db 241 GCCTTGGCCAGCTCGGTGAGCTCTCCGCTGAGACTCACTCAACCGCTGAGCAG 300  
QY 622 CTGGGCTCCGAGACCCGCTTTCTCTGCGGAGGTGATGACAGGCTCTGCGCCCTG 681  
Db 301 CTGGGCTCCGAGACCCGCTTTCTCTGCGGAGGTGATGACAGGCTCTGCGCCCTG 360  
QY 682 GTGCTGAGCTTTAGCGGGAACCCCTGACATGACATGAGCTGAGCTGAGCTGAGCGG 741  
Db 361 GTGCTGAGCTTTAGCGGGAACCCCTGACATGAGCTGAGCTGAGCTGAGCGG 420  
QY 742 CTGGCGCGCGGAGCACTGGAACGTGCGCTCCCGCCGCGCTGAGCGCGCTAC 801  
Db 421 CTGGCGCGCGGAGCACTGGAACGTGCGCTCCCGCCGCGCTGAGCGCGCTAC 480  
QY 802 TTCTGGGCACTGCGCGGAGGCGGCTTCTCTGAGGCGGCTCTCATTTGCGCGCACG 861  
Db 481 TTCTGGGCACTGCGCGGAGGCGGCTTCTCTGAGGCGGCTCTCATTTGCGCGCACG 540  
QY 862 CAGCGGCTCTGGGCTGAGGAAGCGGCGGCGGCGGCTGCGGCGGCGGCTGAG 921  
Db 541 CAGCGGCTCTGGGCTGAGGAAGCGGCGGCGGCGGCTGCGGCGGCGGCTGAG 600  
QY 922 CCGCGGCTTACCATGAGTGGGTCTGAGCAGCGGCTGAGTGGTGGCAACTCTCCGA 981  
Db 601 CCGCGGCTTACCATGAGTGGGTCTGAGCAGCGGCTGAGTGGTGGCAACTCTCCGA 660  
QY 982 GCGCGGCTTACCATGAGTGGGTCTGAGCAGCGGCTGAGTGGTGGCAACTCTCCGA 1041  
Db 661 GCGCGGCTTACCATGAGTGGGTCTGAGCAGCGGCTGAGTGGTGGCAACTCTCCGA 720  
QY 1042 GGTACACCTGATCGGCAACCACTCTGCTGAGGAGGCGGCGGAGTGAAGTACGCG 1101  
Db 721 GGTACACCTGATCGGCAACCACTCTGCTGAGGAGGCGGCGGAGTGAAGTACGCG 780  
QY 1102 GTGCTGGCTTGGCCCATGCTGGGGAACAGCAGTGGCGGAGGCGGCGGCGGCTG 1161  
Db 781 GTGCTGGCTTGGCCCATGCTGGGGAACAGCAGTGGCGGAGGCGGCGGCGGCTG 840  
QY 1162 GACATGCGCGGCTCGGCTGCTGCTGCGGAGGAGTGAAGGAGTGAAGTGAAGTGA 1221  
Db 841 GACATGCGCGGCTCGGCTGCTGCTGCGGAGGAGTGAAGGAGTGAAGTGAAGTGA 900  
QY 1222 GCGGTGACAGGTGACGAGGTGACCGGCACTCAGGAGCTGAGTGGGCTGCGGAGG 1281  
Db 901 GCGGTGACAGGTGACGAGGTGACCGGCACTCAGGAGCTGAGTGGGCTGCGGAGG 960  
QY 1282 CCAGCGGACCCAGTGTGATGTTCCAAATCGATCAACAGCAGGAGATGAGACCTTC 1341

Db 961  
1342 ATCTACCGGATTTGCCAGGCTCCAGGACCACTTCTGCTGGAAGCACTGCTCCCGG 1401  
Db 1021 ATCTACCGGATTTGCCAGGCTCCAGGACCACTTCTGCTGGAAGCACTGCTCCCGG 1080  
QY 1402 GCTGACTATGACCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1461  
Db 1081 GCTGACTATGACCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
QY 1462 GCGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1521  
Db 1141 GCGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
QY 1522 GCGCTGAGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1581  
Db 1201 GCGCTGAGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
QY 1582 GCTGCTTACTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1641  
Db 1261 GCTGCTTACTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
QY 1642 CGCTCCCGCTCAAGCTCAGGCAAGCTCAGTCCAGCAACCAATGAGAGGCGGAGG 1701  
Db 1321 CGCTCCCGCTCAAGCTCAGGCAAGCTCAGTCCAGCAACCAATGAGAGGCGGAGG 1380  
QY 1702 CCGAGGCGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1761  
Db 1381 CCGAGGCGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440  
QY 1762 CTGGGAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1821  
Db 1441 CTGGGAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
QY 1822 AGGCACTCTGTGATGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1881  
Db 1501 AGGCACTCTGTGATGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
QY 1882 GAGCGGCTGGAAGAGTGTGTGA 1908  
Db 1561 GAGCGGCTGGAAGAGTGTGTGA 1587

RESULT 5  
BC023036 2553 bp mRNA linear ROD 30-JUN-2004  
LOCUS  
DEFINITION  
Mus musculus leucine rich repeat and fibronectin type III domain  
containing 4, mRNA (cdna clone MGC:36464 IMAGE:5358628), complete  
cds.  
ACCESSION  
BC023036  
VERSION  
BC023036.1 GI:27696757  
KEYWORDS  
MGC.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 2553)  
Strasberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, K.H., Shemmer, C.F., Bhat, N.K.,  
Altschul, S.F., Zeeberg, B., Buetow, K.R., Schaefer, C.M., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Datchenko, L., Marnett, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Tothyluk, S.,  
Carninci, P., Prange, C., Rana, S.S., Loguigliano, N.A., Peters, G.J.,  
Abramson, R.D., Mollath, S.J., Bosak, S.A., McMan, P.J.,  
McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huliy, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Faney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalka, U., Smalins, D.E., Schermerh, A., Schein, J.E., Jones, S.J. and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H., Kowal, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAX Plate: 54 Row: 1 Column: 5  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23510244.

FEATURES  
source

Location/Qualifiers  
1..2553  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="MGC:36464 IMAGE:5358628"  
/issue\_type="Eye, retina, mouse strain C57BL/6"  
/clone\_lib="NIH MGC\_94"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1..2553  
/gene="Trfn4"  
/note="synonyms: MGC36464, MGC36545"  
/db\_xref="LocusID:225875"  
329..2239  
/db\_xref="MGI:2385612"  
/gene="Trfn4"  
/codon\_start=1  
/product="leucine rich repeat and fibronectin type III domain containing 4"  
/protein\_id="AAH23036.1"  
/db\_xref="GI:27696758"  
/db\_xref="LocusID:225875"  
/db\_xref="MGI:2385612"

CDS

/translation="KAPLLLLIASGAACPVCQNSLSLTLCAHRLFPF  
NNDRLVELRLADNFIALGPPDFRMTGLVDLTLRNAITRGASFDLESLRLH  
LDGNRLVELGSSSLRGPVNLHLSGNQGRAPAFDFLSDLDLVSNNRQV  
PWAGIGMPALHTLNDHNLIDALPGVFAQLSRLDLSRLATLADPLFSGR  
DAASPSPLVTSNGNHCNCELMLRLARPDLETCASPTLAGRYFAVEGEF  
SCPEPLIARTQRLTLEGRATLRCALEDPPVPMHWGPPDRILVNSRMAFPNG  
TEITAGAGDAGAVTCIATNPAGEATARVELVLAHPGNTSAGSGRPSDIAS  
ARTAGAGETLESEPAVOYEVATSGVSWGGRADPVMMQIOYNSSEDTLIR  
IVASSGHTLKLIVGADYDLCLIALSPAGPSDLTATRLCAHSTLPAPLHA  
LOAHVGHFLTAVGVVLAALVFTALLVRRGAGSGLPLKLSHVOSQGTGSP  
MPKSHPRSPRPORSCLDLDTGTGCYARLRGLGAMARRSHSVHGLGACRGGV  
GSAERLREBSVV"

ORIGIN

Query Match

79.9%; Score 1525.4; DB 10; Length 2553;

Best Local Similarity 87.8%; Pred. No. 2.2e-202;  
Matches 1677; Conservative 0; Mismatches 231; Indels 3; Gaps 1;

QY	1	ATGAGCCCCCGGCTCTGCTGCTGTCGACAGTGGAGGGCGCGCTGCGCC	60
DB	329	ATGAGCCCCCGGCTCTGCTGCTGTCGACAGTGGAGGGCGCGCTGCGCC	388
QY	61	TGCGTCTGCGAAGCTGTGCGAGTGTGCGACCCCTGTGCGCCAGGCTGTG	120
DB	389	TGTGTGTCCAGAACTGTGTGCGAGTGTGCGACCCCTGTGCGCCAGGCTGTG	448
QY	121	TTTGTGCGCCCAAGTGGACCGCGGACAGTGGAGCTGCGGCTGTGCAACTTC	180
DB	449	TTTGTGCGCCCAAGTGGACCGCGGACAGTGGAGCTGCGGCTGTGCAACTTC	508
QY	181	CAGGCGCTGGAGCCCCCTGACCTTCCGCAATGACGGGACCTGATGACCTGCT	240
DB	509	CAGGCGCTGGAGCCCCCTGACCTTCCGCAATGACGGGACCTGATGACCTGCT	568
QY	241	CGCAATGCGCATCACCGGATTTGGGACCGCGCTTTGGGGAACCTTGAGAGCTTGC	300
DB	569	CGCAATGCGCATCACCGGATTTGGGACCGCGCTTTGGGGAACCTTGAGAGCTTGC	628
QY	301	CTCAGCTTTGACGGCAAGGCTGTGAGTGTGAGCTGGGACCGGAGGCTCCGGGCTC	360
DB	629	TTGCACTGTGAGTGGCAAGGCTGTGAGTGTGAGCTGGGACCGGAGGCTCCGGGCTC	688
QY	361	AATCGAGACCTGATCTGATCGGCGGCAACAGCTGGGCGGATGCGCGGAGGCTTC	420
DB	669	AATCGAGACCTGATCTGATCGGCGGCAACAGCTGGGCGGATGCGCGGAGGCTTC	748
QY	421	GACGACTTCTAGAGAGCTGTGAGAGCTGTGACCTGTCTCAACAACTTCGGGAGTG	480
DB	749	GACGACTTCTGTGACAGTGTGAGAGCTGTGACCTGTCTCAACAACTTCGGGAGTG	808
QY	481	CCCTGGGCGGATGTGGGCGCATGCTGTGCTGCGACACCTTCAACTGACATACCTT	540
DB	809	CCCTGGGCGGATGTGGGCGCATGCTGTGCTGCGACACCTTCAACTGACATACCTT	868
QY	541	ATTGAGCGACCTGCGGCGGCGGCTTGGCGGAGCTCGGATCGGCTCGGCTGACCTC	600
DB	869	ATTGAGCGACCTGCGGCGGCGGCTTGGCGGAGCTCGGATCGGCTCGGCTGACCTC	928
QY	601	ACCTCCAAACCGGCTGTGAGCAAGCTGTGCTCGGACCCGCTTTCTGTGAGGCTGATGA	660
DB	929	ACCTCCAAACCGGCTGTGAGCAAGCTGTGCTCGGACCCGCTTTCTGTGAGGCTGATGA	988
QY	661	GAGGCTTCTCGGCGGCTGTGCTGCTGAGCTTTAGGGGAAACCCCTGTGACATGTG	720
DB	989	GAGGCTTCTCGGCGGCTGTGCTGCTGAGCTTTAGGGGAAACCCCTGTGACATGTG	1048
QY	721	GAGGCTTCTGAGGCGGCGGCTGTGAGGCGGCGGACGACCTGTGAAAGCTGCGCTCCG	780
DB	1049	GAGGCTTCTGAGGCGGCGGCTGTGAGGCGGCGGACGACCTGTGAAAGCTGCGCTCCG	1108
QY	781	CCCGGCTGTGAGGCGGCGGCTGTGAGGCGGCGGACGAGGCGGAGTTCTTCTGTGAGCG	840
DB	1109	CCCGGCTGTGAGGCGGCGGCTGTGAGGCGGCGGACGAGGCGGAGTTCTTCTGTGAGCG	1168
QY	841	CCCTTATTGCGGCGGCGGCGGCTGTGAGGCGGCGGACGAGGCGGAGTTCTTCTGTGAGCG	900
DB	1169	CCCTTATTGCGGCGGCGGCGGCTGTGAGGCGGCGGACGAGGCGGAGTTCTTCTGTGAGCG	1228
QY	901	CGGTGCGGCGGCGGCGGCTGTGAGGCGGCGGCTGTGAGGCGGCGGAGTTCTTCTGTGAGCG	960
DB	1229	CGGTGCGGCGGCGGCGGCTGTGAGGCGGCGGCTGTGAGGCGGCGGAGTTCTTCTGTGAGCG	1288
QY	961	TTGCTTGGCAACTCTCTCCGAGGCGCGGCTTTTCCCAACCGGACCTTGAAGATTGGGCTG	1020
DB	1289	TTGCTTGGCAACTCTCTCCGAGGCGCGGCTTTTCCCAACCGGACCTTGAAGATTGGGCTG	1348
QY	1021	ACCGGCTGTGAGGCGGCTGTGAGGCGGCTGTGAGGCGGCGGAGTTCTTCTGTGAGCGG	1080

Db 1349 AGCGGCGCTGAGATGAGAGAGCCTATACCTGATTCGACCAACCTCTGTGTAGGCC 1408

Qy 1081 ACAGCCCGAGTGAAGTGGGAGTGTGGCTTGGCCCATGGTGGGAACAGAGGCCGAG 1140

Db 1409 ACAGCCCGAGTGAAGTGGGAGTGTGGCTTGGCCCATGGTGGGAACAGAGGCCGAG 1468

Qy 1141 GGG 1200

Db 1469 GGG 1528

Qy 1201 GGGAGCGCTGAGTGAAGTGGGAGTGTGGCTTGGCCCATGGTGGGAACAGAGGCCGAG 1260

Db 1529 GGAACCTTGAATGATGAACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1588

Qy 1261 GTGAGTGGGGTCCCGGG 1320

Db 1589 GTGAGTGGGGTCCCGGG 1648

Qy 1321 AGCAGCGGAATGAGACCTTCATCTACCGGATTTGTCCAGCCTTCAGCACAACCTTCCTG 1380

Db 1649 AGCAGCGGAATGAGACCTTCATCTACCGGATTTGTCCAGCCTTCAGCACAACCTTCCTG 1708

Qy 1381 CTGAGGACCTTCGCTCCCGGG 1440

Db 1709 CTGAGGACCTTCGCTCCCGGG 1768

Qy 1441 GCTGGGCGCTCTGACCTCAGCGGCACAGCGCTGTGGGGTGTGCCCATTTCTCCAGCGTG 1500

Db 1769 GCTGGGCGCTCTGACCTCAGCGGCACAGCGCTGTGGGGTGTGCCCATTTCTCCAGCGTG 1828

Qy 1501 CCGGCGCTCCGCTGTGGCCAGCGCCCTTCAGAGGCCCACTGTGGGGGGGGGGGGGGGGGG 1560

Db 1829 CCGGCGCTCCGCTGTGGCCAGCGCCCTTCAGAGGCCCACTGTGGGGGGGGGGGGGGGGGG 1888

Qy 1561 GCGGTGGGGGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1620

Db 1889 GCGGTGGGGGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1948

Qy 1621 GCGGTGGGGGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1680

Db 1949 GCGGTGGGGGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 2008

Qy 1681 AATGGAGGCGCCAGCCCAACCCAGAGGCCCAACCCAGAGGCCCAACCCAGAGGCCCAACCCAG 1740

Db 2009 AATGGAGGCGCCAGCCCAACCCAGAGGCCCAACCCAGAGGCCCAACCCAGAGGCCCAACCCAG 2068

Qy 1741 CAGCGAGCTGTCTCTGTGACCTGTGAGATGCG---GGGTGCTAGGTTATGCGAGGCC 1797

Db 2069 CAGCGAGCTGTCTCTGTGACCTGTGAGATGCG---GGGTGCTAGGTTATGCGAGGCC 2128

Qy 1798 CTGGGAGAGCTTGGGCGCGAGCGAGCCCACTGTGATGAGGGGGCTGCTGGGGGAGCGAG 1857

Db 2129 CTGGGAGAGCTTGGGCGCGAGCGAGCCCACTGTGATGAGGGGGCTGCTGGGGGAGCTGGG 2188

Qy 1858 TGCCGGGGGGGTGAGAGGCGCGAGCGCGCTGTGAAGAAGTGTGTGTGA 1908

Db 2189 TGTCCGGGGGGTGGGGGGAGTGTGGAGCGCGCTGTGAAGAAGTGTGTGTGA 2239

RESULT 6  
BC023156 2612 bp mRNA linear ROD 25-MAR-2004  
LOCUS Mus musculus leucine rich repeat and fibronectin type III domain  
DEFINITION containing 4, mRNA (cDNA clone MGC:36545 IMAGE:4951386), complete  
cds.  
ACCESSION BC023156  
VERSION BC023156.1 GI:22137562  
KEYWORDS MGC  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1, (bases 1 to 2612)

# AUTHORS JOURNAL PUBMED REFERENCE AUTHORS TITLE JOURNAL

## REMARK COMMENT

## FEATURES source

gene

CDS

Straussberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G.,  
Klausner, R. D., Collins, F. S., Wagner, L., Shennan, C. M., Schuler, G. D.,  
Altschul, S. F., Zeeberg, B., Bueltow, R. H., Schaefer, C. F., Bae, N. K.,  
Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Hsieh, F.,  
Datchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L.,  
Stapleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L.,  
Scheetz, T. E., Brownstein, M. J., Uddin, T. B., Teshnyuk, S.,  
Carninci, P., Prange, C., Raha, S. S., Loggellano, N. A., Peters, G. J.,  
Abramson, R. D., Mulian, S. J., Bosak, S. A., McEwan, P. J.,  
McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S.,  
Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W.,  
Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A.,  
Fahney, J., Helton, E., Ketterman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y.,  
Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D.,  
Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schultz, J., Myers, R. M.,  
Butterfield, Y. S., Krzywinski, M. I., Skalska, U., Smallus, D. E.,  
Scherer, A., Schein, J. E., Jones, S. J., Jones, S. J., and Marra, M. A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 2612)  
Strausberg, R.  
Direct Submission  
Submitted (04-FEB-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Jeffrey Green M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnar, P. H., Garcia, A. M., Lu, X., Hulyk, S. W., Louie, H.,  
Kowis, C. R., Sneed, A. J., Martin, R. G., Muzny, D. M., Nanavati,  
A. N., Gibbs, R. A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/BLND at: <http://image.llnl.gov>  
Series: IRAX Plate: 60 Row: K Column: 4  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 23510244.  
Location/Qualifiers  
1..2612  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="MGC:36545 IMAGE:4951386"  
/tissue\_type="Mammary tumor, C3(1)-Tag model. Infiltrating  
ductal Carcinoma. 5 month old virgin mouse."  
/clone\_lib="NCI CGAP Mam6"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SF0Rn6"  
1..2612  
/gene="Lrtn4"  
/note="synonyms: MGC36545, MGC36464"  
/db\_xref="locusID:225875"  
/db\_xref="MGI:2385612"  
372..2282  
/gene="Lrtn4"  
/codon\_start=1  
/product="Lrtn4 protein"  
/protein\_id="AAH23156.1"  
/db\_xref="GI:22137565"  
/db\_xref="locusID:225875"







```

Db      2052 AATGATGACACAGCCCATGCCCCAAGAGCAACCCACAGAGAGCCCTCCACCCCTCCT 2111
Qy      1741 CAGGGAGCTGTCTTCTGACCTTGGAGATGCC---GGTGTCTACGGTTATGCCAGGCGC 1797
Db      2112 CAGGGAGCTGTCTTCTGACCTTGGAGATGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2171
Qy      1798 CTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1857
Db      2172 CTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2231
Qy      1858 TGCCGGGGGGGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1908
Db      2232 TGTCGGGGGGGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2282

```

```

RESULT 7
AP000485 139072 bp DNA linear PRI 07-JUN-2002
LOCUS Homo sapiens genomic DNA, chromosome 11, clone: CMB9-21L24, complete
DEFINITION
ACCESSION AP000485
VERSION AP000485.5 GI:21327916
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hattori, M., Toyoda, A., Taylor, T. D., Fujiyama, A., Yada, T.,
Toto, Y., Watanabe, H., and Sakaki, Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (1999)
REFERENCE 2 (bases 1 to 139072)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong, Seog, P.,
Fujiyama, A., Yada, T., Toto, Y., Watanabe, H., and Sakaki, Y.
TITLE Direct Submision
JOURNAL Submitted (13-SEP-1999) Masahisa Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suenho-cho, Tsukuba, Ibaraki, Japan
Tel: 81-45-503-9111, Fax: 81-45-503-9170
URL: http://hgp.gsc.riken.go.jp/
COMMENT On Jun 6, 2002 this sequence version replaced gi:11990449.
FEATURES
source
1..139072
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="CMB9-21L24"

```

```

ORIGIN
Query Match 71.0%; Score 1355; DB 9; Length 139072;
Best Local Similarity 71.8%; Pred. No. 2.7e-179;
Matches 1908; Conservative 0; Mismatches 0; Indels 543; Gaps 1;

```

```

Qy      1 ATGGCCCGCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db      131471 ATGGCCCGCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 131530
Qy      61 TCGCTCTGCAAGACCTGTCGAGTGTGCTGACAGACCCCTCTGTCGCCACCGAGAGCTGTG 120
Db      131531 TCGCTCTGCAAGACCTGTCGAGTGTGCTGACAGACCCCTCTGTCGCCACCGAGAGCTGTG 131590
Qy      122 TTTTGGCCGCCCAAGCTGAGACCGGCGCAGAGTGTGAGCTGTGCTGTGCAAACTTCAATC 180
Db      131591 TTTTGGCCGCCCAAGCTGAGACCGGCGCAGAGTGTGAGCTGTGCTGTGCAAACTTCAATC 131650
Qy      181 CAGGCGCTGGGGCGCCCTGACTTTCGCAACATGACGGGAGCTGATGGAACCTGACACCTGCT 240
Db      131651 CAGGCGCTGGGGCGCCCTGACTTTCGCAACATGACGGGAGCTGATGGAACCTGACACCTGCT 131710

```

```

Qy      241 CGCAATGCCATCAACCGGCAATTTGGGGCCCGGCTTTTGGGGAGCTTCGAGAGACTTGGCTTC 300
Db      131711 CGCAATGCCATCAACCGGCAATTTGGGGCCCGGCTTTTGGGGAGCTTCGAGAGACTTGGCTTC 131770
Qy      301 CTCACCTTGAACGGCAACAGGCTGTGTGAGAGCTGGGCAACGGAGAGCTCCGGGGCCCGGTC 360
Db      131771 CTCACCTTGAACGGCAACAGGCTGTGTGAGAGCTGGGCAACGGAGAGCTCCGGGGCCCGGTC 131830
Qy      361 AATTCGAGACCTTCAATCTTCAAGCTGCAACAGCTGGGCGGCAATGCGCGGAGAGCTTTC 420
Db      131831 AATTCGAGACCTTCAATCTTCAAGCTGCAACAGCTGGGCGGCAATGCGCGGAGAGCTTTC 131890
Qy      421 GACGACTTCTTGAAGAGCTGTGAGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGAGCTGTG 480
Db      131891 GACGACTTCTTGAAGAGCTGTGAGAGAGCTGTGAGAGCTGTGAGAGAGCTGTGAGAGAGCTGTG 131950
Qy      481 CCCTGGGCGGCGCATGCGGCGCATGCTGCTGCTGCAACCTTCAACCTGAGCACTTAACCTT 540
Db      131951 CCCTGGGCGGCGCATGCGGCGCATGCTGCTGCTGCAACCTTCAACCTGAGCACTTAACCTT 132010
Qy      541 ATTGAACGACATGCGCCGAGAGCGCTTGGCCAGCTGTGCAAGCTTCCGCGCTGAGCTTC 600
Db      132011 ATTGAACGACATGCGCCGAGAGCGCTTGGCCAGCTGTGCAAGCTTCCGCGCTGAGCTTC 132070
Qy      601 ACCTCCAGCGGCTGCGCAAGCTGCTCGGAGAGCGGCTTTCTGTGTGGGCGTGAATGCA 660
Db      132071 ACCTCCAGCGGCTGCGCAAGCTGCTCGGAGAGCGGCTTTCTGTGTGGGCGTGAATGCA 132130
Qy      661 GAGGCTCTTCCGCGCCCGCTGTGTGCTGAGCTTTAGGGGAAACCCCTGCACTGCAACTGT 720
Db      132131 GAGGCTCTTCCGCGCCCGCTGTGTGCTGAGCTTTAGGGGAAACCCCTGCACTGCAACTGT 132190
Qy      721 GAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db      132191 GAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 132250
Qy      781 CCCGCGCTGCGCGCGCGCTACTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db      132251 CCCGCGCTGCGCGCGCGCTACTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 132310
Qy      841 CCCCTCATTTGCCCGCCACAGCAGCGCTCTGAGGAGCTGAGAAAGCCAGCGGCGCAAGCTG 900
Db      132311 CCCCTCATTTGCCCGCCACAGCAGCGCTCTGAGGAGCTGAGAAAGCCAGCGGCGCAAGCTG 132370
Qy      901 CGGTGCGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Db      132371 CGGTGCGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 132430
Qy      961 TTGTTTGGCAACTCTCCCGGAGAGCGGAGCTTCCCAACGGGACCTTAGAGATTGGGGTG 1020
Db      132431 TTGTTTGGCAACTCTCCCGGAGAGCGGAGCTTCCCAACGGGACCTTAGAGATTGGGGTG 132490
Qy      1021 ACCGCGCTGGGAGAGCTTGGGGGCTTACACCTGATGCCACCAACCTTGTGTGAGAGCC 1080
Db      132491 ACCGCGCTGGGAGAGCTTGGGGGCTTACACCTGATGCCACCAACCTTGTGTGAGAGCC 132550
Qy      1081 ACAGCCGAGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Db      132551 ACAGCCGAGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 132610
Qy      1141 GGGGGCGCGCCCGGCGCTTCCGACATGCGCGGCTTCCGCTCGAGTGTGTCGAGAGGTGAG 1200
Db      132611 GGGGGCGCGCCCGGCGCTTCCGACATGCGCGGCTTCCGCTCGAGTGTGTCGAGAGGTGAG 132670
Qy      1201 GGGAGCTGTGAGTCTGAGCGAGCGGTGCAAGTGTGAGCGGAGGTGACCGCACTCAAGGGCTG 1260
Db      132671 GGGAGCTGTGAGTCTGAGCGAGCGGTGCAAGTGTGAGCGGAGGTGACCGCACTCAAGGGCTG 132730
Qy      1261 GTGAGCTGGGGGTCGGGGGCGGCGAGCGGAGCCAGAGTGTGAGTGTGCAATCCAGTGAAC 1320
Db      132731 GTGAGCTGGGGGTCGGGGGCGGCGAGCGGAGCCAGAGTGTGAGTGTGCAATCCAGTGAAC 132790
Qy      1321 AGCAGCGAAGATGAGACCTCATTTACC----- 1348

```

```

Db      132791 AGCAGGAAAGATGAGACCTCTACCGGTGAGATGCGTCCCAACCGGCTGCAC 132850
      1349 ----- 1348
Qy      132851 TCCCGGCGCCCTTCTCCGCGCTCTGCTCCCAAGGCTTTGTTCCACCCCTCTCTC 132910
      1349 ----- 1348
Db      132911 TCTGGGGGCTGACACCCCTCCCGACATGCGTGTGACTTTGAGAGAGAGTGGCCT 132970
      1349 ----- 1348
Qy      132971 GGCCTGCGCTGCTGCTGAGGGGCTGAGCTCAAGGGCGGAGAAAGAACCCCAACCC 133030
      1349 ----- 1348
Db      133031 GAGATGTCCAGAGTGTCACTGAGCTGTGCAAGGCTTCCCGGAGAGAGTGGCGTTGGGT 133090
      1349 ----- 1348
Qy      133091 CTGTCTGAAGCAGATGGCCCGCGCGCTTATTCTTCTGCTCTGCTGTGTCATCCCTCC 133150
      1349 ----- 1348
Db      133151 CCACCCATTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 133210
      1349 ----- 1348
Qy      133211 AACTCTTCAGAGCTGTTTCACTTCTCTGAGTGAACCCAGAGCCCGAGTGGCCCGCAGGGGG 133270
      1349 ----- 1348
Db      133271 AAGGGGTGTTGAGAGCTGGAGAGCAGGGAGTGGGGAGTGGGGCTAGACCCACG 133330
      1349 ----- 1348
Qy      133331 CCCGGGCTTGACACACTGCGCTTGCTGAGAGATTGTCCAGAGCTTCAGACCAACACTTC 133390
      1349 ----- 1348
Db      13378 CTGCTAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1437
      13391 CTGCTAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 133450
      1438 GCGGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1497
      13451 GCGGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 133510
      1498 CTGCGGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1557
      133511 CTGCGGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 133570
      1558 GTGGCGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1617
      133571 GTGGCGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 133630
      1618 CCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1677
      133631 CCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 133690
      1678 ACCAATGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1737
      133691 ACCAATGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 133750
      1738 CCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1797
      133751 CCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 133810
      1798 CTGGGAGAGAGCTTGGGCGCGAGCGAGCACTGTGTGATGGGGGCGCTGCTGGGGGAGAGGG 1857
      133811 CTGGGAGAGAGCTTGGGCGCGAGCGAGCACTGTGTGATGGGGGCGCTGCTGGGGGAGAGGG 133870
      1858 TGGCGGGGGGTAGAGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1908

```

```

Db      133871 TGGCGGGGGGTAGAGAGCGAGCGCGCGCGCTGGAAGAGAGTGTGTGA 133921
      182769 bp DNA linear PRI 27-APR-2002
      Homo sapiens genomic DNA, chromosome 11q clone:RP11-15L8, complete
      sequences.
      AP003176
      AP003176
      AP003176.2 GI:20334337
      HTG.
      SOURCE
      ORGANISM
      Homo sapiens (human)
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
      REFERENCE
      1
      Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
      Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
      Homo sapiens genomic DNA
      Published Only in Database (2001)
      2 (bases 1 to 182769)
      Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
      Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
      Direct Submission
      Submitted (09-FEB-2001) Masahira Hattori, The Institute of Physical
      and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan
      1-7-22 Suehiro-chou, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan
      (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
      Tel:81-45-503-9111, Fax:81-45-503-9170)
      On Apr 26, 2002 this sequence version replaced gi:12863263.
      COMMENT
      source
      FEATURES
      1..182769
      Location/Qualifiers
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="11"
      /map="11q"
      /clone="RP11-15L8"
      ORIGIN
      Query Match 71.0%; Score 1355; DB 9; Length 182769;
      Best Local Similarity 77.8%; Pred. No. 2,4e-179; Indels 543; Gaps 1;
      Matches 1908; Conservative 0; Mismatches 0;
      Qy      1 ATGCCCCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
      Db      1845 ATGCCCCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1904
      Qy      61 TGGCTGTGCGAAGCTGTGCGAGTGTGCTGAGACCTCTGTGGCCAGAGGCGCTGCTG 120
      Db      1905 TGGCTGTGCGAAGCTGTGCGAGTGTGCTGAGACCTCTGTGGCCAGAGGCGCTGCTG 1964
      Qy      121 TTTTGGCGCGCAAGTGTGAGCGCGGCGAGTGTGAGCTGGGCTGTGACAACTTCATC 180
      Db      1965 TTTTGGCGCGCAAGTGTGAGCGCGGCGAGTGTGAGCTGGGCTGTGACAACTTCATC 2024
      Qy      181 CAGGCGCTGGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
      Db      2025 CAGGCGCTGGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2084
      Qy      241 CGCAATGCATACACCGCATTTGGGGCGCGCGCTTTGGGGAAGCTGAGAGGCTGCGTTCC 300
      Db      2085 CGCAATGCATACACCGCATTTGGGGCGCGCGCTTTGGGGAAGCTGAGAGGCTGCGTTCC 2144
      Qy      301 CTCCACCTTGAAGCGGCAAGCGCTGTGAGAGCTGGGAGCGGGAGCTCCGGGGCGCCCGTC 360
      Db      2145 CTCCACCTTGAAGCGGCAAGCGCTGTGAGAGCTGGGAGCGGGAGCTCCGGGGCGCCCGTC 2204
      Qy      361 AATCTGACAGCACTTCTCTGACCGGCAACAGCTGTGGCGCATGCGCGCGGAGGCTTTC 420
      Db      2205 AATCTGACAGCACTTCTCTGACCGGCAACAGCTGTGGCGCATGCGCGCGGAGGCTTTC 2264
      Qy      421 GACGACTTCTTAGAGAGCTGTGAGAGCTGTGACTGTCTTCAACAACACTCCGGCAGGTG 480

```





```

QY 721 GAGCTGCTGTGCTGCTGCGGCGCTGAGCGCGGACGACCTGAAACGTTGCTCCCGC 780
Db 5047 GAGCTGCTGTGCTGCTGCGGCGCTGAGCGCGGACGACCTGAAACGTTGCTCCCGC 4988
QY 781 CCGGCGCTGCGCGGCGCTGAGCTTCTGTGCGAGTTCCTCTGTGAGCCG 840
Db 4987 CCGGCGCTGCGCGGCGCTGAGCTTCTGTGCGAGTTCCTCTGTGAGCCG 4928
QY 841 CCCCTCATTTGCGCGGCGGACGAGCGGCTCTGGGTGCTGGAAGGCCGCGGCGCACGCTG 900
Db 4927 CCCCTCATTTGCGCGGCGGACGAGCGGCTCTGGGTGCTGGAAGGCCGCGGCGCACGCTG 4868
QY 901 CCGTGGCGGCGGCGCTGAGTGAACCCGCGGCTTACCACTGCACTGAGTGGTCTCTGAGCAGC 960
Db 4867 CCGTGGCGGCGGCGCTGAGTGAACCCGCGGCTTACCACTGCACTGAGTGGTCTCTGAGCAGC 4808
QY 961 TTGGTTGGCAACTCTCTCCGAGCCGCGGCTTTCCCGCAACGCGGACCTTAGAGATTGGGGTG 1020
Db 4807 TTGGTTGGCAACTCTCTCCGAGCCGCGGCTTTCCCGCAACGCGGACCTTAGAGATTGGGGTG 4748
QY 1021 ACCGCGCTGCGGCGGACGCTGCGGCGGCTAGACCTGATGCGCACCAACCTGCTGTGAGAGCC 1080
Db 4747 ACCGCGCTGCGGCGGACGCTGCGGCGGCTAGACCTGATGCGCACCAACCTGCTGTGAGAGCC 4688
QY 1081 ACAAGCCGAGTAGAACTGCGGCGGCTGCTGCGCTTGCCTGATGAGTGGGAAACAGCAGTGGAG 1140
Db 4687 ACAAGCCGAGTAGAACTGCGGCGGCTGCTGCGCTTGCCTGATGAGTGGGAAACAGCAGTGGAG 4628
QY 1141 GGGGGCGCGCGCGCGCGCGCTTGGGACATGCGCGCTTCCGCTGACATGCTGCGGAGGTGAG 1200
Db 4627 GGGGGCGCGCGCGCGCGCGCTTGGGACATGCGCGCTTCCGCTGACATGCTGCGGAGGTGAG 4569
QY 1201 GGGAGCGCTGAGTCTGAGCGGCGGCTGAGCTGAGTGAACCGGACCTGACAGGCGCTG 1260
Db 4568 GGGAGCGCTGAGTCTGAGCGGCGGCTGAGCTGAGTGAACCGGACCTGACAGGCGCTG 4509
QY 1261 GTAGAGCTGCGGCGGCGCGGCGGCGGACGCGACGAGTGTGATGTTCCAAATCCAGTACAAC 1320
Db 4508 GTAGAGCTGCGGCGGCGCGGCGGCGGACGCGACGAGTGTGATGTTCCAAATCCAGTACAAC 4449
QY 1321 AGAGAGGAATAGAACCTTATCTATCCGAGTGTGCCAGCTTCCAGCCACCACTTCTG 1380
Db 4448 AGAGAGGAATAGAACCTTATCTATCCGAGTGTGCCAGCTTCCAGCCACCACTTCTG 4389
QY 1381 CTGAGACGACCTCTCCCGGCGGCTGACTATGACCTGCTGCGGCGCTTGTACCGGCGC 1440
Db 4388 CCGGCGATCTTCTCTCCGCGCTTGTGCTGCGGCGGCTTGTCTTCCACCCCTCTCTCT 4329
QY 1441 GCTGGCGCTCTGACTGACGCGGCGGCGGCTGCTGGCGTGT 1482
Db 4328 CTTGGGCGCTGACACCCCTCTCCCGACCATGCTGCTGGAATCT 4287

```

REFERENCE  
AUTHORS

RESULT 10  
 BC007718 1844 bp mRNA linear PRI 12-FEB-2004  
 LOCUS Homo sapiens leucine rich repeat and fibronectin type III domain  
 DEFINITION containing 4, mRNA (cDNA clone IMAGE:4110922), partial cds.  
 ACCESSION BC007718  
 VERSION BC007718.2 GI:33873739  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1844)  
 Strussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshynski, S.,

## CDS

## gene

FEATURES  
source

Carninci, P., Prange, C., Raha, S.S., Loguettano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Ketterman, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, D., Myers, R.M.,  
 Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smallos, D.E.,  
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 2 (bases 1 to 1844)  
 Strausberg, R.  
 Direct Submission  
 Submitted (11-MAY-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 On Aug 19, 2003 this sequence version replaced gi:14712783.  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kong, P., Laric, P., Legapi, R.,  
 Maduro, Q.L., Masillo, C., Maskell, B., Mastrian, S.D., McLooney, J.C.,  
 McDowell, J., Pearson, R., Stantipod, S., Thomas, P.J., Touchman, J.W.,  
 Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 17 Row: e Column: 2  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 22095365.  
 Location/Qualifiers  
 1..1844  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4110922"  
 /tissue\_type="muscle, rhabdomyosarcoma"  
 /clone\_lib="NIH MGC 17"  
 /lab\_host="DH10B-R"  
 /note="vector: POTB7"  
 1..>1844  
 /gene="LRFN4"  
 /note="synonym: MGC3103"  
 /db\_xref="LOCUSID:78999"  
 221..>1844  
 /gene="LRFN4"  
 /codon\_start=1  
 /product="LRFN4 protein"  
 /protein\_id="AAH07718.1"  
 /db\_xref="GI:33873740"  
 /db\_xref="LOCUSID:78999"  
 /translation="MAPLLILLASGAACPDPVCONLSSTLCAHRLGVPP  
 NVDRVLEIRADNFIQALGPDPFNMGVLDLTLSRAIIRIGARFEDLSLSLH  
 LDGNLVEIGTOSLGPVNLQHLISGNQILAGADPDFESLIEDLSYNNLRQV  
 PWAGIGAMPALHTLNLIDALPPGAFAQGGRSDLSRLDLSNRATLADPLFSRGR







Db 1330 GGGGGCGCCGCTGGCTGGGA 1352

RESULT 12  
AC119332

LOCUS  
DEFINITION  
AC119332 238317 bp DNA linear HTG 15-NOV-2002  
Rattus norvegicus clone CH230-257B16, WORKING DRAFT SEQUENCE, 2  
unordered pieces.

ACCESSION  
AC119332 GI:25009638  
HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.

KEYWORDS  
Rattus norvegicus  
Rattus norvegicus  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

ORGANISM  
Rattus.

REFERENCE  
1 (bases 1 to 238317)  
Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bismato, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bunay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Drape, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, J., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hollander, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, D., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Louleaged, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, B., Mamhney, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munitasa, M., Murphy, M., Nait, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackelme, O., Okunolu, G., Olarinpoogoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Polindexter, A., Popovic, D., Plums, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Register, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shattman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G., and Gibbs, R. A.

TITLE  
JOURNAL  
Unpublished  
2 (bases 1 to 238317)  
Worley, K. C.  
Direct Submission  
Submitted (26-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One

# REFERENCE

Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 238317)  
Rat Genome Sequencing Consortium.  
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 15, 2002 this sequence version replaced gi:23810333.

## COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

### ----- Genome Center

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GNUMD  
Center clone name: CH230-257B16

### ----- Summary Statistics

Assembly program: Phrap; version 0.990329  
Consensus quality: 214183 bases at least Q40  
Consensus quality: 215804 bases at least Q30  
Consensus quality: 216645 bases at least Q20  
Estimated insert size: 217293; sum-of-contigs estimation  
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

## FEATURES

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 236862: contig of 236862 bp in length  
\* 236863 236962: gap of unknown length  
\* 236963 238317: contig of 1355 bp in length.

### Location/Qualifiers

1. 238317  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-257B16"  
1. 1411  
/note="wgs end\_extension  
clone\_end:T7"  
23470. 24380  
/note="clone\_boundary  
clone\_end:T7  
site:  
end\_sequence:B2142517"  
23470. 235378  
/note="clone\_boundary  
clone\_end:Sp6  
site:  
end\_sequence:B2142518"  
235804. 236862  
/note="wgs end\_extension  
clone\_end:Sp6"

## ORIGIN



Db	211551	ACAGCGCTGAGATGCAAGAGACCTATACCTGATCTGACCAACCTGCGGTGAGGCC	2116110
QY	1081	ACAGCCCGAGGTGAATCTGCGGGTGCCTGGGCTTGGCCCATGTGGGAAACAGACTGTCCGAG	1140
Db	211611	ACAGCCCGAGGTGAGCTCCGGGTAATGSGCTTGGCCCATGTGGGAAACAGACTGTCTGAG	2116707
QY	1141	GGGGGCGCGCCCGGGGCCCTCGGACATCGCGGCTCCGCTCGCACTCTCGCCGAGGGTGA	1200
Db	211671	GGGGGCGCGCTCGGGGCCCTCGGACATGTGCTGCTCTGCTAGAACTGTCTGTGAAGCGAG	211730
QY	1201	GGGACGCTGAGTCTGAGCCAGCCCGTGCAGGTGACGAGAGTGAACGCCACCTCAGGCTTG	1260
Db	211731	GGAACTTTAAGATCTGAACCAACCGCGCAAGTGAACAGAGGTGACTCCACTCTGGACTG	211790
QY	1261	GTGACCTGGGGGTCCCGGGGGCGGCGACGCCGCCAGTGGATGTTCCAAATCCAGTAAAC	1320
Db	211791	GTGACTGGGGGCCCGAGGGCGGCCAGCTGAACCAAGTGGATGTTCCAAATCCAGTAAAC	2118506
QY	1321	AGCAGCGAAGATGAGACCTCTCATCTAC-----	1348
Db	211851	AGCAGCGAAGATGAGAACCTCTCATCTACCGGTGAGAGATGGGCTTGACACTCCCTTCAGC	2119110
QY	1349	-----	1348
Db	211911	CTGCTTGTCTACTGTGCTCCCTTCTGTGACCTGTACCCGCCAAGTTCTGCTGTT	2119707
QY	1349	-----	1348
Db	211971	CCCTTCACAGGCTGTGCTGCTCTCTCTACTTAACTTTGAGTATGAGCTTTGGGACGA	212030
QY	1349	-----	1348
Db	212031	GAAACGTGAGCTGCTTTTAAAGGGGTCTCAGTGTAGGTGTATGTGCAACAGCTTCTCTA	212090
QY	1349	-----	1348
Db	212091	CCATCAGTGTATGACAGCTNN	212150
QY	1349	-----	1348
Db	212151	NN	212210
QY	1349	-----	1348
Db	212211	CCTGAAAGAAAGGAATGGGGGTGTAGCCCAACCCCTGGGGCTCTGACCACTGTGCCTTGGC	2122707
QY	1349	-----	1348
Db	212271	TACAGGATGTACACAGCTCCAGCCAAACACTTCTGTGTAACACTGTGTTCTTGTTGCT	212330
QY	1405	GACTATGACCTCTGCTGCTGCTGCTGTGCAACGGCGCGCTGGGGCCCTGTACACTCAAGGCG	1464
Db	212311	GACTACGACCTCTGCTGCTGCTGCTGTGCACTGCTGCTGGGCCCTCCACCTCAAGGCG	212390
QY	1465	ACCAAGCTGCTGGGCTGTGCCATTCTTCAAGCTGCGGAGCTGTGGCCCTGTGCAAGCGC	1524
Db	212391	ACTAGACTGCTGGGCTGTGCCACTTCTTACCTTACAGGCAACCCCTCTGTGCCATGCG	212450
QY	1525	CTGACAGGCCACGTGTGGGCGGACCTTGAACGTTGCGCGGTGGGGGTGTGCTGTGAGCT	1584
Db	212451	CTACAGGCGCCATGTGCTGGGCGGAGCCTGACTGTGGCAGTGGAGGCGTCTTGTGAGCT	212510
QY	1585	GCGTTACTGTGCTTCACTGTGGCCTTGTGTGGTTCGGGGCCGGGGAGCTGGAGATGGCGC	1644
Db	212511	GCGTTACTGTGCTTCACTGTGGCCTTGTGTGGTTCGGGGCCGGGGAGCTGGAGATGGCGC	212570
QY	1645	CTCCCGCTCAAGCTCAGCGACAGTCCAGTCCAGACCAATGAGAGGCCCGACCCACACCC	1704
Db	212571	CTCCCACTCAAACTCAGCACTGTCCAAATCCAGACCAATGTGTGACTAGCCCATGCCC	212630
QY	1705	AAGGCCCAACCGCGCGAGGCCCGCGCGCGCGCGCAGCGACGTGCTCTTGTGACCTG	1764

```

Db      212631 AAGAGCCACCCACGACGAGCCCTCCACCCCTGCCAGCGAGCTGCTCCTGGAAGCTA 212690
QY      1765 GGAGATGCC---GGGTGCTACGGTATATGCCAGGCGCTG3GAGAGAGCTTG3GGCCCGGACG 1821
Db      212691 GGAAGACCCGGTGGATGCTACGGGATATGCCAGGCGCTG3GAGAGAGCTTG3GGCCCGG 212750
QY      1822 AGGCACTGTGTGATG3GGGCGCTGCTG3GGGAGAGGAGTGGCCGGGGGTAGAGGACGCGC 1881
Db      212751 AGGCACTGTGTACATG3GGGCGGCTGCTG3GGAGCTGGGTGTG3GGGTATG3GGGCGGAGTGG 212810
QY      1882 GAGCGGCTG3GAGAGAGTGTGTTGTA 1908
Db      212811 GAGCGGCTG3GAGAGAGCGTGTGTA 212837

RESULT 13
AX358338
LOCUS      AX358338      1653 bp      DNA      linear      PAT 13-FEB-2002
DEFINITION Sequence 9 from Patent WO0202604.
ACCESSION AX358338
VERSION    AX358338.1 GI:18674974
KEYWORDS
SOURCE     synthetic construct
ORGANISM   synthetic construct
            artificial sequences.
REFERENCE  1
            Thayer,E.C., Sheppard,P.O. and Presnell,S.R.
            Leucine-rich repeat proteins, zlr7, zlr8 and zlr9
            Patent: WO 0202604-A 9 10-JAN-2002;
            Zymogenetics, Inc. (US)
FEATURES
            source
            1..1653
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="Degenerate polynucleotide sequence"

ORIGIN
Query Match      57.7%; Score 1101.8; DB 6; Length 1653;
Best Local Similarity 55.7%; Pred. No.1.4e-143;
Matches 920; Conservative 349; Mismatches 381; Indels 2; Gaps 2;

QY      1 ATGGCCCGCGCGCTCCTGCTGCTGCTGCTGCGCAGTGAAGCGGCGCGCTGCCGCTGCC 60
Db      1 ATGGCNCNCNCNTYNTYNTYNTYNTYNTYNTGNCMNGCNCNCNCNTGYCCNYTNCN 60
QY      61 TGGGTGCTGCAGAACTGTGCCAGTGGCTGCTGAGACCTCTGTGCCACGAGGCGCTGCTG 120
Db      61 TGGTGTGTGCARAAYTNNMNGARMSNYTNNMNACTYNTGTGNCNCAYMNGNNTYNTN 120
QY      121 TTGTGCGCGCCCAAGCTGAGCCGGCGCAGTGAAGCTGGCGCTGGCTGCAACTTCATC 180
Db      121 TTGTGNCNCNCNAAYTNGAYTNGMNGACGTNGATYNTNGTNCNGAYATYTAH 180
QY      181 CAGGCGCTGGGGCCCCCTGACTTCCGCAATGAGCGGAGCTGTGACCTGACACTGTCT 240
Db      181 CARGNYTNGNCNCNCNGAYTTCMGNAAYATGACNGNYTNGTNGAYTTNA.CNYTNNMS 240
QY      241 CGCAATGCCATCACCCGCAATTTGGGGCCCGGCTTTGGGGAGCTCGAGAGCCTGCTTC 300
Db      241 MGNAAGCNAATHACNNMGNATHGNGCNCMNGCNTTGTGGNGAYTTNGARMSNYTNNMNS 300
QY      301 CTCACCTTGAAGCGCAACAGAGCTGTGAGAGCTGGGACCGGAGCCTCCGGGCGCCGCTC 360
Db      301 YTMCAATYTNAGAYGNAAYMGNATYNTNGATYTNMGNAACNGMNSNYTNNMNGACNGTN 360
QY      361 AATCTGACAGCACTCATCTCAAGCGCAACAGAGCTGGGCGGATCGCGCGGAGGCTTC 420
Db      361 AAYTNCARCAAYTATYTNATHYTNMNGNAAYCARATNGMNGATYTNMNGCNCNGNCNTTY 420
QY      421 GAGGACTTCTTAAGAGCGCTGAGAGAGCTGAGAGCTGTGCTCAACAACACTCGGAGAGTG 480
Db      421 GAGGATYTYTNGARMSNYTNGARAYTNGAYTTMSTAYTAAYAYTNNMNGACNGTN 480

```

```

QY      481 CCTGG3CCGCGAGCTG3GCGCATGCTGCTGCTGACACACCTGACACTGACATTAACCTT 540
Db      481 CCNTGGGCGNGAHTHNGNCNATGCGNCNGNTYNTCAAGNYTAAYTTNAGAYCAAYAYTN 540
QY      541 ATTGAAGCACTGCCCCAGAGCGCTTCGCCAGCTTGCTGCTGAGCTCTCCGCTGACCTTC 600
Db      541 ATGAGAYGNCYTNCCNCNGNCNGNCNTTYGNCNCARYTNNGNCNCARYTNMNSNGNYTN 600
QY      601 ACCTCCACCGGCTGCGCACGCTGCTCCGACCCGCTTCTCTGCTGGGGGTGATGCA 660
Db      601 ACNWSNAAYMGNATYTNCAACNYTNGCNCNGAYCCNYNTYTNMNSNGMNGAYGCA 660
QY      661 GAGGCGCTTCCGCGCCCTGCTGCTGAGCTTGAAGGGAAACCCCTGACCTGACCACTGT 720
Db      661 GARGCNSNCNCNCNCNTYNTNGTNTYTNMNTTYMNSNGNAAYCCNYTNCAYGYAAYTY 720
QY      721 GAGTGTGCTGTGCTGCGGCGGCTGCGGCGGCGGAGACCTGGAACGTCGCCCTCCCG 780
Db      721 GARITYNTVTGGYTNMNGMNTYTNMNGMNCNGAYGAYTYTGARACNTGYCNCMSNCN 780
QY      781 CCGGCGCTGCGCGCGGCGCTACTTCTGGGCGAGTCCCGAGGGAGGTTCCTGTGAGCG 840
Db      781 CNGGNYTNGCNGMNGNTATYTTGTGGGNGTNCNGARGNGARTTYSNTGYARCCN 840
QY      841 CCCCTCATTTCCCGCCACAGCGAGCGCTCTGTGGTGTGGAAGGCCAGCGGCGCACGCTG 900
Db      841 CCNYNTATHTGCMNGMCAVACNARNGNTYNTGGTNTYTNARGNGNARBNGNCNACNYTN 900
QY      901 CGGTGCGGCGCGCTGGGTGACCCCGGCGCTACCATGACTGTGGGTGGTCTTACGACCG 960
Db      901 MNTGTMGNGCNYTNGNAYCCNGCNCNACNATGCAATGTGGTNGCNCNGAYGMYGN 960
QY      961 TTGGTGGCACTCTCCCGGAGCGCGGCTTCCCAAGGGAGCCTTAGAGATTGGGGTG 1020
Db      961 YTNGTNGNAAYMWSMNSMNGNCNMNGNCNTTYCCNAAYGNAACNTNGARATHGNGTN 1020
QY      1021 ACCGCGCTGGGAGCGCTG3G3GCTACACCTGATGCTGCCACCAACCTTGTGAGGCC 1080
Db      1021 ACNNGGNCNGNGAGAYGNCNGNGNTYTA.CNTGYATHGCAACNAAYCCNCNGNGNGARGCN 1080
QY      1081 ACAGCCCGAGTGAACCTGCGGCTGCTGCTGCTGCCATGTGTGGGAACAGCACTGCCGAG 1140
Db      1081 ACNGCMNGTNGATYTNMNGNTYTNNGCNYTNCNCAAGGNGNAAYMNSMNGNCNGAR 1140
QY      1141 GGGGCGCGCGCGGCGCTCGGACATGCGCGCTCGCTCGGACTGCTGCGGAGGTGAG 1200
Db      1141 GGNNGMNGCNGNGCNCNGNAGYATHGNGCNSMNGMNAACNGCNGNGARGNGAR 1200
QY      1201 GGGAGCGCTGAGTCTGAGCCAGCCGCTGACAGTGAAGAGGTGAACCGCACTCAGGGCTG 1260
Db      1201 GGNACNYTNGARMSNGARCCNCNGTNCARGTNACNGAGTNAACNGACNMSNGNYTN 1260
QY      1261 GTGAGCTGGGCTCCGGGCGGCGGACCGACCCAGCTGTGATGTTCCAAATCCAGTCAAC 1320
Db      1261 GTMWSNTGGGNCNCNGMNGCNCNGCNGAYCCNGTNTGATGTYCARATHCARITYAAY 1320
QY      1321 AGCAGGAGATGAGACCCCTCACTACCGGATTTGCTCCAGCCTCCAGCACCACTTCGT 1380
Db      1321 WSNWSNGARGAAGARCNTNTATHTAYMNAATHGNCNGCMWSMNSNCAVAYTYTYTN 1380
QY      1381 CTGAAGCACTGCTCCCGCGGCTGACTTGAACCTTGTGCTGCTGCTGTCACCGGCGC 1440
Db      1381 YTNAAACAYTNTGTCNCNGNGNGAYTA.YGAYTYTNTYTNNGCNYTNNMNCNGCN 1440
QY      1441 GCTGGGCGCTCTGAACCTCAAGGCAACAGCTGTGGGCTGTGCCATTTCTCAAGCTG 1500
Db      1441 GNGGNCNMWSNGAYTYTNA.CNMGNAACNMNYNTYTNMNGTGTGNCAYTYTYMNAACNYTN 1500
QY      1501 CCGGCGCTGCGCGCTGAGCAGCGCTGAGGACCACTGTCTGGGCGGAGCCCTGACCGTG 1560
Db      1501 CNGCNCMSNCNTYNTGTCA.YGNCYTNARGCNCA.YGTNTYTNNGNGMNAACNTNACNGTN 1560

```

[illegible]

```

JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 269255)
AUTHORS      Worley,K.C.
TITLE        Direct Submission
JOURNAL      Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
              3 (bases 1 to 269255)
              Rat Genome Sequencing Consortium.
              Direct Submission
              Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
COMMENT       On May 13, 2003 this sequence version replaced gi:23121847.
              The sequence in this assembly is a combination of BAC based reads
              and whole genome shotgun sequencing reads assembled using Atlas
              (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
              in the feature table below represents a scaffold in the Atlas
              assembly (a 'contig-scaffold'). Within each contig-scaffold,
              individual sequence contigs are ordered and oriented, and separated
              by sized gaps filled with Ns to the estimated size. The sequence
              may extend beyond the ends of the clone and there may be sequence
              contigs within a contig-scaffold that consist entirely of whole
              genome shotgun sequence reads. Both end sequences and whole genome
              shotgun sequence only contigs will be indicated in the feature
              table.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GJKL
Center clone name: CH230-102H22
----- Summary Statistics -----
Assembly program: Atlas 3.0;
Consensus quality: 240505 bases at least Q40
Consensus quality: 242683 bases at least Q30
Consensus quality: 244329 bases at least Q20
Estimated insert size: 251974; sum-of-contigs estimation
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 266736: contig of 266736 bp in length
*   266737 266835: gap of unknown length
*   266837 267942: contig of 1106 bp in length
*   267943 268042: gap of unknown length
*   268043 269255: contig of 1213 bp in length.
* Location/Qualifiers
*     1..269255
*       /organism="Rattus norvegicus"
*       /mol_type="genomic DNA"
*       /db_xref="taxon:10116"
*       /clone="CH230-102H22"
*       /note="wgs_contig"

FEATURES             misc_feature
source               1..1418
                    /note="wgs_contig"

ORIGIN
Query Match          56.8%; Score 1083.8; DB 2; Length 269255;
Best Local Similarity 86.5%; Pred. NO. 8,9e-142;
Matches 1196; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

1 ATGAGCCCGCGCTCCTGCATGTCAGTCAAGAGGGCCGTGCGCGCTGCC

```

```

Db      30653 ATGGCCCGCGCTCCATTTGCTGTGCTGCGCAGTGAAGCGCTGTCGCCGCTGCC
QY      61 TGGCTGCGAGAACTGTCGAGTGTGCTGACAGCCTGTGCGCCACGAGGCGCTG
Db      30713 TGTGTGCGAGAACTGTGTGAGTGTCTGAGCACTTGTGCGCCAGAGGCGCTG
QY      121 TTTGTGCGCGCCAACTGAGACCGGCGCAGAGTGAAGTGTGCTGAGCACTTGTG
Db      30773 TTTGTGCGAACCAAGCTGAGCAGAGGCGCAGATTGAAGTGTGCTGAGCACTTGTG
QY      181 CAGGCGCTGAGGCGCGCTTGAATGCAATGACGAGGCTGGTGAACCTTAATGTCT
Db      30833 CAGGCGCTGAGGCGCGCTTGAATGCAATGACGAGGCTGGTGAACCTTAATGTCT
QY      241 CGGAATGCGATCAACCGGCAATTTGGGCGCGCGCTTTGGGGAAGCTGAGAGCTGCTTC
Db      30893 CGAAATGCGATCAACCGGCAATTTGGGCGCGCGCTTTGGGGAAGCTGAGAGCTGCTTC
QY      301 CTCACCTTGAACGAGCAACAGAGCTGTGAGAGTGTGAGCAGCGGAGCTTCGCGGCGCGCTTC
Db      30953 TTGCACTTGAGTGTGCAACAGGCTGTGTGGAATGTGGGAGAGCAAGCTTACGCGGCGCTTC
QY      361 AATCTGAGAGCCTCACTCTCAAGCGGCAACAGCTGAGCGCGATGCGCGCGAGCTTC
Db      31013 AACCTTCACAGCAGCTCTATCTTGTGTGCAATCAAGCTGAGCGCGATGCGCTTCGCGGCGCTTC
QY      421 GAGGACTTCTGAGAGCGCTGAGAGCACTGTGCTGCTCAACAAGAACTCCGCGGAGTG
Db      31073 GAGGACTTCTGAGAGCGCTGAGAGCACTGTGCTGCTCAACAAGAACTCCGCGGAGTG
QY      481 CCTTGAGCGCGGATCGGCGCGCAATGCTGCTGCTGCAACCTTCAACCTTGAGACCTT
Db      31133 CCTTGAGCGCGGATCGGCGCGCAATGCTGCTGCTGCAACCTTCAACCTTGAGACCTT
QY      541 ATTGAGCACTGCGCGCGCGCGCTTGTGCGCGAGCTGTGAGCTTCCGCGCTGAGCTTC
Db      31193 ATTGAGCGCTGCGCGCGCGCGCGCTTGTGCGCGAGCTGTGAGCTTCCGCGCTGAGCTTC
QY      601 ACCTCCAGCGCTGCGCGCGCGCGCTTGTGCGCGAGCTGTGAGCTTCCGCGCTGAGCTTC
Db      31253 ACCTCCAGCGCTGCGCGCGCGCGCTTGTGCGCGAGCTGTGAGCTTCCGCGCTGAGCTTC
QY      661 GAGGCGCTTCCGCGCGCGCGCGCTTGTGCTGAGCTTGTGAGGGAACCCCTGAGCTGCACTGT
Db      31313 GAGGCGCTTCCGCGCGCGCGCGCTTGTGCTGAGCTTGTGAGGGAACCCCTGAGCTGCACTGT
QY      721 GAGCTGCTGTGCTGCGCGCGCTGTGCGCGCGCGCGAGCACTGTGAAACGTGCGCTTC
Db      31373 GAGCTGCTGTGCTGCGCGCGCGCTGTGCGCGCGCGCGAGCACTGTGAAACGTGCGCTTC
QY      781 CCGGCGCTGCGCGCGCGCGCTTGTGCGCGAGTGTGCGCGAGGCGAGTTCCTGTGAGCGG
Db      31433 CCAACCTTGTGCGCGCGCGCTTGTGCGCGAGTGTGCGCGAGGCGAGTTCCTGTGAGCGG
QY      841 CCGCTCATTTGCGCGCGCGCGCTTGTGCGCGAGTGTGCGCGAGGCGAGGCGCGAGCTG
Db      31493 CCACTGATTTGCGCGCGCGCGCTTGTGCGCGAGTGTGCGCGAGGCGAGGCGAGCTG
QY      901 CCGTGTGCGCGCGCTTGTGCGCGAGTGTGCGCGAGTGTGCGCGAGGCGAGGCGAGCTG
Db      31553 CCGTGTGCGCGCGCTTGTGCGCGAGTGTGCGCGAGTGTGCGCGAGGCGAGGCGAGCTG
QY      961 TTGGTGTGCGCACTTCCGCGAGCGCGGCTTTCGCAAGGAGGCGAGCTTGTGAGGCGG
Db      31613 CTGGTTTGTGCACTTCTGTGAGCTGTGAGCTTTCGCAAGGAGGCGAGCTTGTGAGGCGG
QY      1021 ACCGCGCTGTGAGGAGCGCTGTGAGGAGCTACCTGAGTGTGAGCAACCACTGTGTGAGGCG
Db      31673 ACAGGCGCTGTGAGATGCAAGAGGCTTAACTGTGATTTGCAACCACTGTGTGAGGCG
QY      1081 ACAGCGCGATGAACTGTGCGGCTGTGCGCTTTCGCGCATGTGTGGAACAGCACTGTGCGAG

```

```

Db      31733 ACAGCGCGATGAGAGCTCCGCGGTACTGCGCTTTCGCCCATGTGTGGAACCAAGTGTGAG
QY      1141 GGGGCGCGCGCGCGCGCGCTTGTGAGATGCGCGCGCTGCGCTGAGCTGTGCGAGGCTGAG
Db      31793 GGGGCGCGCTTGTGCGCGCGCGCGCTTGTGAGATGCGCGCGCTGCGCTGAGCTGTGAG
QY      1201 GGGAGCTGTGAGTGTGAGCGAGCGCGCTGTGAGAGTGTGAGAGCGCGCGCGCTTGTGAG
Db      31853 GGAACTTTAGATCTGAAACCAAGCGGAGAGTGTGAGAGTGTGAGAGCGCGCGCTTGTGAG
QY      1261 GTGAGCTGTGAGGCTTCCGCGCGCGCGCGCGCTTGTGAGATGTTCCAAATCCAGTGTGAG
Db      31913 GTGAGCTGTGAGGCTTCCGCGCGCGCGCGCGCTTGTGAGATGTTCCAAATCCAGTGTGAG
QY      1321 AGCAGCGAGATGAGAACCTCTATCAAGGATTTGTCCAGCTCCAGCGACCACTTCCG
Db      31973 AGCAGCGAGATGAGAACCTCTATCAAGGATTTGTCCAGCTCCAGCGACCACTTCCG
QY      1381 CTG 1383
Db      32033 CTG 32035

RESULT 15
AC128739 195342 bp DNA linear ROD 11-NOV-2003
LOCUS Mus musculus BAC clone RP24-168G14 from chromosome 19, complete
DEFINITION
AC128739
VERSION AC128739.4 GI:28475716
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE
AUTHORS Waligorski,J., Haglund,R. and Bielicki,L.
TITLE The sequence of Mus musculus BAC clone RP24-168G14
JOURNAL Unpublished (2001)
AUTHORS Wilson,R.
TITLE Sequencing of Mus musculus
JOURNAL Unpublished (2001)
REFERENCE
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submision
JOURNAL Submitted (23-JUL-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 195342)
REFERENCE
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submision
JOURNAL Submitted (19-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 195342)
REFERENCE
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submision
JOURNAL Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 195342)
REFERENCE
AUTHORS Wilson,R.
TITLE Direct Submision
JOURNAL Submitted (11-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Feb 23, 2003 this sequence version replaced gi:28416270.
COMMENT
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
-----
Summary Statistics
-----
Center project name: M_BB0168614

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:  
The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:  
This sequence is the entire insert of the clone. This clone is overlapped by AC124347.

#### FEATURES

##### source

1.195342  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="19"  
/map="19"  
/clone="RP24-168G14"  
/clone\_1fb="RPCI-24"  
99..242  
/rpt\_family="Alu"  
332..529  
/rpt\_family="B2"  
576..723  
/rpt\_family="Alu"  
583..731  
/rpt\_family="B4"  
732..870  
/rpt\_family="Alu"  
871..945  
/rpt\_family="B4"  
945..1065  
/rpt\_family="Alu"  
1073..1331  
/rpt\_family="L1"  
1454..1800  
/rpt\_family="MaLR"  
2820..2962  
/rpt\_family="Alu"  
3003..3068  
/rpt\_family="MER53"  
3086..3281  
/rpt\_family="B2"  
3557..3720  
/rpt\_family="B4"  
3923..3982  
/rpt\_family="L1"  
3982..4077  
/rpt\_family="B4"  
4226..4357  
/rpt\_family="B4"  
5765..5912  
/rpt\_family="Alu"  
5937..6108  
/rpt\_family="B4"

repeat\_region 6558..6905  
/rpt\_family="MaLR"  
repeat\_region 7443..7522  
/rpt\_family="L1"  
repeat\_region 7828..7976  
/rpt\_family="Alu"  
8124..8278  
/rpt\_family="Alu"  
8582..8759  
/rpt\_family="B2"  
8895..9034  
/rpt\_family="B4"  
9214..9317  
/rpt\_family="L1"  
9380..9525  
/rpt\_family="Alu"  
9938..10114  
/rpt\_family="B4"  
10167..10282  
/rpt\_family="Alu"  
10348..10479  
/rpt\_family="Alu"  
10543..10732  
/rpt\_family="B2"  
11623..11746  
/rpt\_family="B4"  
11990..11983  
/rpt\_family="L1"  
12002..13314  
/rpt\_family="L1"  
13479..13559  
/rpt\_family="L1"  
13710..14395  
/rpt\_family="RMR19B"  
14403..14493  
/rpt\_family="B4"  
15359..15506  
/rpt\_family="Alu"  
15712..16025  
/rpt\_family="MaLR"  
16270..16397  
/rpt\_family="RMR15"  
16501..16624  
/rpt\_family="Alu"  
16705..16836  
/rpt\_family="ERVK"  
16919..16952  
/rpt\_family="ERVK"  
16933..17169  
/rpt\_family="ERVK"  
17511..17576  
/rpt\_family="Alu"  
17651..17779  
/rpt\_family="Alu"  
17780..17951  
/rpt\_family="MER1\_type"  
17952..17982  
/rpt\_family="B4"  
18633..19157  
/rpt\_family="L1"  
19511..19639  
/rpt\_family="B4"  
19888..20036  
/rpt\_family="L1"  
20081..20463  
/rpt\_family="MaLR"  
20467..20608  
/rpt\_family="L1"  
20650..20826  
/rpt\_family="B2"  
21346..21464  
/rpt\_family="B4"  
21531..21680

```

/rpt_family="Alu"
repeat_region 23326. 23512
/rpt_family="MAIR"
repeat_region 23553. 23715
/rpt_family="B4"
repeat_region 23916. 24055

```

Query Match	56.7%;	Score 1082.2;	DB 10;	Length 195342;
Best Local Similarity	86.4%;	Pred. No. 1.6e-141;		
Matches 1195; Conservative	0;	Mismatches 188;	Indels 0;	Gaps 0;

[illegible]

Db	81573	CGGTGCAAGGACCCTTGATGTAACCTGTACTACCAATGCACTGGATGGCCCTGATGACAAG	81632
Qy	961	TTGGTTGGCAACTCTCTCCGAGGCCCGGCTTTCCCAACGGGACCTTAGAGATTGGGGTG	1020
Db	81633	CTGTGTTGGCAACTCTCTCGAGGCTGGGGCTTTTCCCAATGGGACCTTAGAGATTGGGGTG	81692
Qy	1021	ACGGGCGCTGGGGACGCTGAGGGGCTACACTGATGCGACCAACCTGCTGGTAGGCC	1080
Db	81693	ACGGGCGCTGGAGATGACAGAGGCTTATCTGCAATTGGCAACCAACCTGCTGGTAGGCC	81753
Qy	1081	ACAGCCGAGTAGAATGCGGGGTGCTGGCCTTTGCCCATGGTGGGAACAGCATGCCGAG	1140
Db	81753	ACAGCCGAGTAGAGCTTCGGGTACTGGGCTTGCCCATGGTGGGAACCAAGTGTGAG	81812
Qy	1141	GGGGGCGGCCCGGGGCGCTCGGACATCGCGCGCTCGGCTCGGACTGCTGCGAGGGGTGAG	1200
Db	81813	GGGGGCGGCTCTGGGCGCTTGGACATTGCTGCTTCTGCTAGAACTGCTGCGAAGCGAG	81872
Qy	1201	GGGACGCTGGAGTCTTAGAGCCAGGCCGTGACAGGTGACGGAGGTGACCGCCACTCAGGGCTG	1260
Db	81873	GGAACTTTAGATCTGATAACACAGACAGTGAAGTGAACGGAGGTGACTGCCACTCTGGCCCTG	81932
Qy	1261	GTAAGCTGGGCTCCCGGGGCGGCGGACCGACCAAGATGTGATGTTCCTCAATCCAGTACAAC	1320
Db	81933	GTAAGCTGGGCGCTGGGGCGGCGGCTGACCAAGATGTGATGTTCCTCAATCCAGTACAAC	81992
Qy	1321	AGCAGGAGAGATAGAACCTCATCTAACGGATTTGTCCACGCTCCAGCCACCACTTCTGTG	1380
Db	81993	AGCAGGAGAGATAGAACCTCATCTAACGGATGTGAGATGTGGCGCTTGCAACACGCGCTTCAAC	82052
Qy	1381	CTG 1383	
Db	82053	CTG 82055	

Search completed: November 20, 2004, 12:18:45  
Job time : 12934 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 17, 2004, 22:32:38 ; Search time 197 Seconds

(without alignments)  
1854.634 Million cell updates/sec

Title: US-10-071-879-10

Perfect score: 3336

Sequence: 1 MAPPILLILLIAGAAACPLP.....LGACRGVGSARLEESVY 635

Scoring table: BLASTSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3336	100.0	635	2 Q6PUG9	Q6PUG9 homo sapien
2	3336	100.0	635	2 AAH15581	Aah15581 homo sapi
3	3327.5	96.7	636	2 Q80XU8	Q80XU8 mus musculu
4	3327.5	96.7	636	2 Q8K3C4	Q8K3C4 mus musculu
5	3327.5	96.7	636	2 BAC38259	Bac38259 m 16 days
6	2784	83.5	528	2 AAH27475	Aah27475 homo sapi
7	2302.5	69.0	541	2 Q6PK41	Q6PK41 homo sapien
8	2302.5	69.0	541	2 AAH07718	Aah07718 homo sapi
9	2013	60.3	460	2 AAH71866	Aah71866 homo sapi
10	2013	60.3	460	2 Q6IPL6	Q6IPL6 homo sapien
11	1733	51.9	329	2 Q713C2	Q713C2 homo sapien
12	1733	51.9	329	2 AAH14040	Aah14040 homo sapi
13	1720	51.6	327	2 Q9BWD0	Q9BWD0 homo sapien
14	1720	51.6	327	2 AAH00207	Aah00207 homo sapi
15	1662.5	49.8	324	2 Q8N644	Q8N644 homo sapien
16	1662.5	49.8	324	2 Q9ULH4	Q9ULH4 homo sapien
17	1657.5	49.7	789	2 Q9BE71	Q9BE71 macaca fasc
18	1644	49.3	833	2 BAC31891	Bac31891 mus muscu
19	1644	49.3	833	2 Q80T99	Q80T99 mus musculu
20	1630	48.9	788	2 Q9CYK3	Q9CYK3 mus musculu
21	1556	46.6	628	2 Q9BTN0	Q9BTN0 homo sapien
22	1556	46.6	628	2 BADI8740	Badi8740 homo sapi
23	1546.5	46.4	627	2 QADY10	Qady10 homo sapien
24	1546.5	46.4	627	2 AAQ88494	Aaq88494 homo sapi
25	1539.5	46.1	626	2 Q8BLY3	Q8BLY3 mus musculu
26	1539.5	46.1	626	2 AAH66999	Aah66999 mus muscu
27	1486	44.3	719	1 LRF5 HUMAN	LRF5 homo sapien
28	1479	39.9	700	2 Q9P244	Q9P244 mus musculu
29	1331.5	38.8	584	2 Q6PKX3	Q6PKX3 homo sapien
30	1293.5	38.8	584	2 AAH56798	Aah56798 brachydan
31	1293.5	38.8	584	2	

32	1223	36.7	492	2 Q99KT6	Q99KT6 mus musculu
33	1139	34.1	463	2 Q8CIV9	Q8CIV9 mus musculu
34	884	26.5	542	2 Q8TBS9	Q8TBS9 homo sapien
35	618.5	18.5	450	2 Q96C50	Q96C50 homo sapien
36	499	15.0	593	2 Q6UY18	Q6UY18 homo sapien
37	499	15.0	593	2 AAQ88482	Aaq88482 homo sapi
38	495	14.8	618	2 BAC30759	Bac30759 m adult m
39	482.5	14.5	709	2 AAH60263	Aah60263 mus muscu
40	456.5	13.7	730	2 Q6PMP6	Q6PMP6 mus musculu
41	456.5	13.7	730	2 AAH56458	Aah56458 mus muscu
42	455.5	13.7	730	2 Q6US92	Q6US92 mus musculu
43	455.5	13.7	730	2 AAQ74241	Aaq74241 mus muscu
44	433.5	13.0	713	2 Q8N182	Q8N182 homo sapien
45	433.5	13.0	713	2 AAH68541	Aah68541 homo sapi

## ALIGNMENTS

RESULT 1  
Q6PUG9 PRELIMINARY; PRT; 635 AA.  
AC Q6PUG9;  
DT 05-JUL-2004 (TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
DE 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
DE Leucine rich repeat and fibronectin type III domain containing 4.  
CN Name-LRRN4;  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schier G.D.,  
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko U., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettlemen M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywicki M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Warr M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RA Strausberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC015581; AAH15581.2; -  
DR InterPro; IPR003961; FN III.  
DR InterPro; IPR008957; FN\_III-like.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR003598; Ig C2.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR003591; LRR\_typ.  
DR Pfam; PF00041; fn3; 1.  
DR Pfam; PF00047; Ig; 1.  
DR Pfam; PF00560; LRR; 7.



Query Match	100.0%;	Score 3336;	DB 2;	Length 635;
Best Local Similarity	100.0%;	Pred. No. 8.3e-200;		
Matches 635;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
DR PRINTS;PRO0019; LEURICHRPT.				
DR SMART; SMO0060; FN3; 1.				
DR SMART; SMO0409; IG; 1.				
DR SMART; SMO0408; IGC2; 1.				
DR SMART; SMO0082; LRCT; 1.				
DR SMART; SMO0369; LRR_TYP; 6.				
DR PROSITE; PS00853; FN3; 1.				
DR PROSITE; PS00835; IG_LIKE; 1.				
DR SEQUENCE 635 AA; 66859 MW; 1366645A17A92EE CRC64;				
QY 1 MAPPLILLLLSGAACPPLPCVQCNLSESLSTLCAHRLFLPVPVNDRTVRLADNFI				60
Db 1 MAPPLILLLLLSGAACPPLPCVQCNLSESLSTLCAHRLFLPVPVNDRTVRLADNFI				60
QY 61 QALGPDRNNMTGLVDLTLSRNAITRIGARAFGDELSRLHLDGNRLVLELGTSLRGV				120
Db 61 QALGPDRNNMTGLVDLTLSRNAITRIGARAFGDELSRLHLDGNRLVLELGTSLRGV				120
QY 121 NLQHLILSGNQGRIPAPAFDFLESLBEDLSTYNNLRQVPAGIGAMPALHTLNDHL				180
Db 121 NLQHLILSGNQGRIPAPAFDFLESLBEDLSTYNNLRQVPAGIGAMPALHTLNDHL				180
QY 181 IDALPPGAPAFQGLSLRLDLSNRALATAPDLFSGRGDAESPAPVLVLSFGNPLHCNC				240
Db 181 IDALPPGAPAFQGLSLRLDLSNRALATAPDLFSGRGDAESPAPVLVLSFGNPLHCNC				240
QY 241 ELMRLRLARPDDLETCASSPGLAGRYFWAVVEGSPCEPPLIARTHTQRLWLEGORATL				300
Db 241 ELMRLRLARPDDLETCASSPGLAGRYFWAVVEGSPCEPPLIARTHTQRLWLEGORATL				300
QY 301 RCRAIGDDPAPTMHWGPDPDLVLGNSSSRAAPFNGTLEIVTGADGAGYTCAIATNPAGEA				360
Db 301 RCRAIGDDPAPTMHWGPDPDLVLGNSSSRAAPFNGTLEIVTGADGAGYTCAIATNPAGEA				360
QY 361 TARVELRLVALPHGNSSSAEBGRPGPSDIASARTAAEGEGTLESHPAVQTEVTAISGL				420
Db 361 TARVELRLVALPHGNSSSAEBGRPGPSDIASARTAAEGEGTLESHPAVQTEVTAISGL				420
QY 421 VSMWGRPADPVMWFQIOYNSSEDEFLYRIYPASSHFLHVLVPGAYDCLIALSPA				480
Db 421 VSMWGRPADPVMWFQIOYNSSEDEFLYRIYPASSHFLHVLVPGAYDCLIALSPA				480
QY 481 AGPSDLTATRLLGCAHFTFLPASPLCHALQAVHLGGTLTVAGVVLVAALLFTVALIVR				540
Db 481 AGPSDLTATRLLGCAHFTFLPASPLCHALQAVHLGGTLTVAGVVLVAALLFTVALIVR				540
QY 541 GRGAANGRLPLKLSHVQSGTNGGSPPTPAHPRRSPPPPPSCSIDLDAGCYGARLL				600
Db 541 GRGAANGRLPLKLSHVQSGTNGGSPPTPAHPRRSPPPPPSCSIDLDAGCYGARLL				600
QY 601 GGAMARRSHSVHGGLLGAGCRGVGSAEKLSESVV 635				
Db 601 GGAMARRSHSVHGGLLGAGCRGVGSAEKLSESVV 635				
RESULT 2				
AAH15581 PRELIMINARY; PRT; 635 AA.				
AC AAH15581;				
DT 02-MAR-2004 (Tremblrel. 27, Created)				
DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)				
DT 02-MAR-2004 (Tremblrel. 27, Last annotation update)				
DE Leucine rich repeat and fibronectin type III domain containing 4.				
GN LRPY4.				
OS Homo sapiens (Human).				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.				
OK NCBI_TaxId=9606;				
RN [1]				

Query	Match	100.0%	Score 3336	DB 2	Length 635
Query	Beat Local Similarity	100.0%	Pred. No. 8.3e-200		
Matches	635	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
1	MAPPLILLILLASGAACACPLPCVCQNLSSLSLTLCARHGLTFPPVAVNDRTVELRLADNFI	60			
1	MAPPLILLILLASGAACACPLPCVCQNLSSLSLTLCARHGLTFPPVAVNDRTVELRLADNFI	60			
61	QALGPDRFNNMTGLVDLTLSRNAITRIGARAFGDLESLSLHDGNRLVELCTGSLRGPV	120			
61	QALGPDRFNNMTGLVDLTLSRNAITRIGARAFGDLESLSLHDGNRLVELCTGSLRGPV	120			
61	QALGPDRFNNMTGLVDLTLSRNAITRIGARAFGDLESLSLHDGNRLVELCTGSLRGPV	120			
121	NIQHILISGNOLGRIPAGAFDFLESLEDDILSYNNLRQVPAGIGAMPALHTLINDHL	180			
121	NIQHILISGNOLGRIPAGAFDFLESLEDDILSYNNLRQVPAGIGAMPALHTLINDHL	180			
121	NIQHILISGNOLGRIPAGAFDFLESLEDDILSYNNLRQVPAGIGAMPALHTLINDHL	180			
181	IDALPPGAFQGLSRLDILTSNRATLAPDPLFSRGRFAESPAVLVLSFGNPLHNC	240			
181	IDALPPGAFQGLSRLDILTSNRATLAPDPLFSRGRFAESPAVLVLSFGNPLHNC	240			
181	IDALPPGAFQGLSRLDILTSNRATLAPDPLFSRGRFAESPAVLVLSFGNPLHNC	240			
241	ELIMRLRLARPDDETCASSPPLAAGRYTVAVBERGESCPLIARHTQMLNMLEGRATL	300			
241	ELIMRLRLARPDDETCASSPPLAAGRYTVAVBERGESCPLIARHTQMLNMLEGRATL	300			
241	ELIMRLRLARPDDETCASSPPLAAGRYTVAVBERGESCPLIARHTQMLNMLEGRATL	300			
301	RCRALGDPAPMTGWGPPDDRVLVGNSSRARAFPNGTLEIGVTGAGDAGGYTCIATNPAGRA	360			
301	RCRALGDPAPMTGWGPPDDRVLVGNSSRARAFPNGTLEIGVTGAGDAGGYTCIATNPAGRA	360			
301	RCRALGDPAPMTGWGPPDDRVLVGNSSRARAFPNGTLEIGVTGAGDAGGYTCIATNPAGRA	360			
361	TARVELRLVALPHGNSSSABGGRPGPSDIAASARTPAEGGTLESEPAVQVTBVTATSGI	420			
361	TARVELRLVALPHGNSSSABGGRPGPSDIAASARTPAEGGTLESEPAVQVTBVTATSGI	420			
361	TARVELRLVALPHGNSSSABGGRPGPSDIAASARTPAEGGTLESEPAVQVTBVTATSGI	420			
421	VSMWGGRRADPVMWQIOYNSSEDETLIRIYPASSHHFLKHLVPAGAYDCLALASPA	480			
421	VSMWGGRRADPVMWQIOYNSSEDETLIRIYPASSHHFLKHLVPAGAYDCLALASPA	480			
421	VSMWGGRRADPVMWQIOYNSSEDETLIRIYPASSHHFLKHLVPAGAYDCLALASPA	480			
481	AGPSDLTATRLTLCGAFHFTLPASPLCHALQAAHVIGGTLTVAGVGVVAALLVFTVALLVR	540			
481	AGPSDLTATRLTLCGAFHFTLPASPLCHALQAAHVIGGTLTVAGVGVVAALLVFTVALLVR	540			
481	AGPSDLTATRLTLCGAFHFTLPASPLCHALQAAHVIGGTLTVAGVGVVAALLVFTVALLVR	540			
541	GRGANGRLPLKLSHVQSTNGSGEPTPEKAHPRRPPRQSSGLDGLDACCYGARRL	600			
541	GRGANGRLPLKLSHVQSTNGSGEPTPEKAHPRRPPRQSSGLDGLDACCYGARRL	600			
541	GRGANGRLPLKLSHVQSTNGSGEPTPEKAHPRRPPRQSSGLDGLDACCYGARRL	600			

Db 541 GRGANGRLPLKLSHVOSQNGBSPPTPKHPPRSPRQSCSLDLDGAGCYGARRL 600  
 QY 601 GGAMARRSHVHGGLGAGRGVGGSAERLEESV 635  
 Db 601 GGAMARRSHVHGGLGAGRGVGGSAERLEESV 635  
 RESULT 3  
 Q80XU8 PRELIMINARY; PRT; 636 AA.  
 ID Q80XU8  
 AC Q80XU8  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Leucine rich repeat and fibronectin type III domain containing 4.  
 GN Name=lrifn4;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye;  
 RA Strausberg R.;  
 RL Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC023036; AAH23036.1; -  
 DR HSSP: Q9BZR6; 1P8T.  
 DR MGD: MGI:2385612; Lrfin4.  
 DR InterPro: IPR003961; FN\_III.  
 DR InterPro: IPR008957; FN\_III-like.  
 DR InterPro: IPR007110; IG-like.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR00047; Ig\_1.  
 DR Pfam: PF00047; Ig\_1.  
 DR Pfam: PF00560; LRR\_7.  
 DR PRINTS: PRO0019; LEURICRPT.  
 DR PROSITE: PS00853; FN3\_1.  
 DR PROSITE: PS00853; IG\_LIKE; 1.  
 SQ SEQUENCE 636 AA; 67251 MW; 1EE86E96C88BA91 CRC64;

Query Match 96.7%; Score 3227.5; DB 2; Length 636;  
 Best Local Similarity 96.7%; Pred. No. 4,9e-193;  
 Matches 615; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

QY 1 MAPPULLLLASGAACPLPCVCQNTSESLTICARGLLFPVNDRTVELRLADNFI 60  
 Db 1 MAPPULLLLASGAACPLPCVCQNTSESLTICARGLLFPVNDRTVELRLADNFI 60

QY 61 QALGPPDFNNMTGLVDLTLSRNAITRIGARFGLDLSRLHDGRLVELGTSLRGPV 120  
 Db 61 QALGPPDFNNMTGLVDLTLSRNAITRIGARFGLDLSRLHDGRLVELGTSLRGPV 120  
 QY 121 NLQHLILSGNQLRIAPGAFDDFLESLEDLSYNNLRQYPMWAGIGAMPALHTLNDHNL 180  
 Db 121 NLQHLILSGNQLRIAPGAFDDFLESLEDLSYNNLRQYPMWAGIGAMPALHTLNDHNL 180  
 QY 181 IDALPPGAFAOLGOLSRDLTSLRLATLAPDPLFSRGRABEAPALVYSPGNPLHCNC 240  
 Db 181 IDALPPGAFAOLGOLSRDLTSLRLATLAPDPLFSRGRABEAPALVYSPGNPLHCNC 240  
 QY 241 ELIMLRRLARPDDLETCAAPGLAGRYFMAVEGEESCEPPLIARTORLMTVEGORATL 300  
 Db 241 ELIMLRRLARPDDLETCAAPGLAGRYFMAVEGEESCEPPLIARTORLMTVEGORATL 300  
 QY 301 RCRALEDPAPTMHWGPDDRLVGNSSRADAPFNGTLEIGVTAGAGAGYTCIATNPAGEA 360  
 Db 301 RCRALEDPAPTMHWGPDDRLVGNSSRADAPFNGTLEIGVTAGAGYTCIATNPAGEA 360  
 QY 361 TARVELRVLALPHGSSAEGRPGSDIAASARTAEGBGLESEPAVQVTEVTATSGI 420  
 Db 361 TARVELRVLALPHGSSAEGRPGSDIAASARTAEGBGLESEPAVQVTEVTATSGI 420  
 QY 421 VSMGPRPADPYWMFQIQNSSDEDTLIRIYPASSHFLKHLVAGADYDCLIALSPA 480  
 Db 421 VSMGPRPADPYWMFQIQNSSDEDTLIRIYPASSHFLKHLVAGADYDCLIALSPA 480  
 QY 481 AGPSDITATRLIGCAHFSTLPASPLCHALQAHVLTGLTVAVGVVLAALVETVALVR 540  
 Db 481 AGPSDITATRLIGCAHFSTLPASPLCHALQAHVLTGLTVAVGVVLAALVETVALVR 540  
 QY 541 GKGAGNGRLPLKLSHVOSQNGSPPTPKHPPRSPRQSCSLDLDGAGCYGAR 599  
 Db 541 GKGAGNGRLPLKLSHVOSQNGSPPTPKHPPRSPRQSCSLDLDGAGCYGAR 599  
 QY 600 LGGAMARRSHVHGGLGAGRGVGGSAERLEESV 635  
 Db 600 LGGAMARRSHVHGGLGAGRGVGGSAERLEESV 635

RESULT 4  
 Q8K3C4 PRELIMINARY; PRT; 636 AA.  
 ID Q8K3C4  
 AC Q8K3C4  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Lrfin4 protein.  
 GN Name=lrifn4;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzyzinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Maira M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
 RA Strauberg R.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC023156; AAR23156.1; -  
 DR HSSP; Q9BZR6; 1P8T.  
 DR MGI; MGI:2385612; Lrfn4.  
 DR InterPro; IPR003961; FN\_1.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000483; LRR\_Cterm.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR InterPro; IPR003591; LRR\_Typ.  
 DR Pfam; PF00041; fn3; 1.  
 DR Pfam; PF00047; Ig; 1.  
 DR Pfam; PF00560; LRR; 7.  
 DR PRINTS; PR00019; LEUR1CHRP1.  
 DR SMART; SM00060; FN3; 1.  
 DR SMART; SM00082; LRCT; 1.  
 DR SMART; SM00013; LRRNT; 1.  
 DR SMART; SM00369; LRR\_Typ; 2.  
 DR SMART; PS50853; FN3; 1.  
 DR PROSITE; PS50835; IG\_Like; 1.  
 DR PROSITE; PS50835; IG\_Like; 1.  
 SQ SEQUENCE 636 AA; 67264 MW; 9CE86E8A3981A884 CRC64;  
 Query Match 96.7%; Score 3227.5; DB 2; Length 636;  
 Best Local Similarity 96.7%; Pred. No. 4.9e-193;  
 Matches 615; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

QY 1 MAPPLILLILASGAACPLPCVCONLSSTLCANHGGLFVPPNVRRITVELRIANFI 60  
 DB 1 MAPPLILLILASGAACPLPCVCONLSSTLCANHGGLFVPPNVRRITVELRIANFI 60  
 QY 61 QALGPPFRMTGIVDLTSLRNATIRIGARFGDLESLRSLHLDGNRLVELIGTSLRGPV 120  
 DB 61 QALGPPFRMTGIVDLTSLRNATIRIGARFGDLESLRSLHLDGNRLVELIGTSLRGPV 120  
 QY 121 NLQHLITSSQGLRIAGARDPFLIESIEDLSTNNNRQVPMAGIGAMPALHTLNLDHNL 180  
 DB 121 NLQHLITSSQGLRIAGARDPFLIESIEDLSTNNNRQVPMAGIGAMPALHTLNLDHNL 180  
 QY 121 NLQHLITSSQGLRIAGARDPFLIESIEDLSTNNNRQVPMAGIGAMPALHTLNLDHNL 180  
 DB 121 NLQHLITSSQGLRIAGARDPFLIESIEDLSTNNNRQVPMAGIGAMPALHTLNLDHNL 180  
 QY 181 IDALPGAPFQGLSLDLTSLNRLATLAPPLPSRGRDAEASAPAVLSFSGNPLHCNC 240  
 DB 181 IDALPGAPFQGLSLDLTSLNRLATLAPPLPSRGRDAEASAPAVLSFSGNPLHCNC 240  
 QY 241 ELIMLRRLARPDDLETGASPPGLAGRYFWAVBEGFSCPEPLIARHTQRLMWLEGQRATL 300  
 DB 241 ELIMLRRLARPDDLETGASPPGLAGRYFWAVBEGFSCPEPLIARHTQRLMWLEGQRATL 300  
 QY 241 ELIMLRRLARPDDLETGASPPGLAGRYFWAVBEGFSCPEPLIARHTQRLMWLEGQRATL 300  
 DB 241 ELIMLRRLARPDDLETGASPPGLAGRYFWAVBEGFSCPEPLIARHTQRLMWLEGQRATL 300  
 QY 301 RCRAIGPPAPMTMHWGDDRLVGNSSRAAPFNPTLEIGTGAGDAGGTCIANPGEA 360  
 DB 301 RCRAIGPPAPMTMHWGDDRLVGNSSRAAPFNPTLEIGTGAGDAGGTCIANPGEA 360  
 QY 361 TARVELRVLALPHGNGSSAEGRPGSPDIAASARTAAEGEGTLESEPAVQTEVTATNSGL 420  
 DB 361 TARVELRVLALPHGNGSSAEGRPGSPDIAASARTAAEGEGTLESEPAVQTEVTATNSGL 420  
 QY 421 VSMGCPGAPADPVMMFOIQYNSSSEDETLIYRIVPASHHFLKHLVPADYDLCLALSPA 480  
 DB 421 VSMGCPGAPADPVMMFOIQYNSSSEDETLIYRIVPASHHFLKHLVPADYDLCLALSPA 480  
 QY 481 AGPSDLTATLTLGCAHSTLPASPLCHALQAHVIGTLTAAGVGVVAALIVFTVALIVR 540  
 DB 481 AGPSDLTATLTLGCAHSTLPASPLCHALQAHVIGTLTAAGVGVVAALIVFTVALIVR 540

QY 541 GRGANGRLPLKLSHVOSQTNGGSPPTPKAHPSPRPPRQRCSDLDG-AGCYGARR 599  
 DB 541 GRGANGRLPLKLSHVOSQTNGGSPPTPKAHPSPRPPRQRCSDLDG-AGCYGARR 600  
 QY 600 LGGAMARRSHSVHGLLGAGCRGVGSGARLEESV 635  
 DB 601 LGGAMARRSHSVHGLLGAGCRGVGSGARLEESV 636  
 RESULT 5  
 BAC38259  
 ID BAC38259 PRELIMINARY; PRT; 636 AA.  
 AC BAC38259;  
 DT 14-APR-2004 (TEMBLrel. 27, Created)  
 DT 14-APR-2004 (TEMBLrel. 27, Last sequence update)  
 DE 16 days embryo head cDNA, RIKEN full-length enriched library,  
 DE clone:CI30040J16 product:hypothetical Immunoglobulin and major  
 DE histocompatibility complex domain/leucine-rich repeat/fibronectin type  
 DE III domain/ Cysteine-rich flanking region, C-terminal containing  
 DE protein, full insert sequence.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=2049374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Kitunai T., Tashiro H., Itoh M.,  
 RA Kono H., Akiyama J., Nishi K., Kitunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaishiki K.,  
 RA Fujiwara S., Inoue K., Togawa Y., Itawa M., Ohara E., Watanabe M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuda S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hatanaka T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiroaka I., Hirose T.,  
 RA Hori F., Imocant K., Ichi Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Komoto H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numataki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Shibata K., Shinagawa A., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Sakai K., Sakazume N., Sano H.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeuchi Y., Tanaka T.,  
 RA Tomaru A., Toyota T., Yasunishi A., Yamamoto S., Hayashizaki Y.,  
 RA Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK081560; BAC38259.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 636 AA; 67251 MW; 1EE86E96CB88BA91 CRC64;  
 Query Match 96.7%; Score 3227.5; DB 2; Length 636;  
 Best Local Similarity 96.7%; Pred. No. 4,9e-193;  
 Matches 615; Conservative 9; Mismatches 11; Indels 1; Gaps 1;  
 QY 1 MAPPRLILLIAGSAAACPLPCVCONLSESLTCAHGLFVPPNDRRTVELRLADNFI 60  
 DB 1 MAPPRLILLIAGSAAACPLPCVCONLSESLTCAHGLFVPPNDRRTVELRLADNFI 60  
 QY 61 QALGPPDFRMTGLVLTLSRNAITRIGAPAFGLDLSRLHLDGNRLVELGTSLRGPV 120  
 DB 61 QALGPPDFRMTGLVLTLSRNAITRIGAPAFGLDLSRLHLDGNRLVELGTSLRGPV 120  
 QY 121 NIQHLISGNQGRIRIAPAFDDLESLDLSYNNLRQVPMAGTAMPALHTLNDHNI 180  
 DB 121 NIQHLISGNQGRIRIAPAFDDLESLDLSYNNLRQVPMAGTAMPALHTLNDHNI 180  
 QY 181 IDALPPGAFQOLGSLRLDLTSNRLATLAPDPLFSRGRDAEAPAPLVISFGNPLHCNC 240  
 DB 181 IDALPPGAFQOLGSLRLDLTSNRLATLAPDPLFSRGRDAEAPAPLVISFGNPLHCNC 240  
 QY 241 ELIMRLRLARPDDLETCASPPGLAGRYFMAVPEGEFSCPEPLIARHTORLWLEGORATL 300  
 DB 241 ELIMRLRLARPDDLETCASPPGLAGRYFMAVPEGEFSCPEPLIARHTORLWLEGORATL 300  
 QY 301 RCRAALGDPAPTMHWVPPDRLVNSSRAAFPNGLTIEGYTGAGDAGYTCTATNPAGA 360  
 DB 301 RCRAALGDPAPTMHWVPPDRLVNSSRAAFPNGLTIEGYTGAGDAGYTCTATNPAGA 360  
 QY 361 TRAVELRVIALPHGNSAEGGRPGPSDIAASARTAEGBGTLESPPAYQVTEVATSG 420  
 DB 361 TRAVELRVIALPHGNSAEGGRPGPSDIAASARTAEGBGTLESPPAYQVTEVATSG 420  
 QY 421 VSMGGRPADPVMFQIYNSSEDETLIYRIYVASSHFLKHLVPGADYDCLIALSPA 480  
 DB 421 VSMGGRPADPVMFQIYNSSEDETLIYRIYVASSHFLKHLVPGADYDCLIALSPA 480  
 QY 481 AGPSDITATRLIGCAHFTLPAIPCHALQAVHGLGTLTVAAVGVVALLVFTVALVR 540  
 DB 481 AGPSDITATRLIGCAHFTLPAIPCHALQAVHGLGTLTVAAVGVVALLVFTVALVR 540  
 QY 541 GRGAGNGRLPLKLSHVQSQTNCGSPPTPKAHPRPSPRPORSCSLDLD AGCYGYAR 599  
 DB 541 GRGAGNGRLPLKLSHVQSQTNCGSPPTPKAHPRPSPRPORSCSLDLD AGCYGYAR 599  
 QY 600 LGGAMARRSHSVHGLLGAGCGVGSAAELEESVY 635  
 DB 600 LGGAMARRSHSVHGLLGAGCGVGSAAELEESVY 635  
 QY 601 LGGAMARRSHSVHGLLGAGCGVGSAAELEESVY 636  
 DB 601 LGGAMARRSHSVHGLLGAGCGVGSAAELEESVY 636

## RESULT 6

AAH27475 PRELIMINARY; PRT; 528 AA.  
 ID AAH27475  
 AC AAH27475;  
 DT 01-JUN-2004 (Tremblrel. 27, Created)  
 DT 01-JUN-2004 (Tremblrel. 27, Last sequence update)  
 DT 01-JUN-2004 (Tremblrel. 27, Last annotation update)

DE LRFN4 protein (Fragment).  
 GN LRFN4.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung.  
 RX MEDLINE=2238257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang Y., Hsieh F.,  
 RA Diatchenko L., Matusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Mada A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shychenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalski U., Smalls D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Matra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung.  
 RA Strausberg R.,  
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC027475; AAH27475.2; -  
 FT NON TER 1  
 SQ SEQUENCE 528 AA; 55315 MW; E5B57A25FF22F2B2 CRC64;  
 Query Match 83.5%; Score 2784; DB 2; Length 528;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-165;  
 Matches 527; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 108 IVELGTSLRGVNIQHLISGNQGRIRIAPAFDDLESLDLSYNNLRQVPMAGTGA 167  
 DB 1 IVELGTSLRGVNIQHLISGNQGRIRIAPAFDDLESLDLSYNNLRQVPMAGTGA 167  
 QY 168 MPALHTLNDHNLIDALPPGAFQOLGSLRLDLTSNRLATLAPDPLFSRGRDAEAPAPLV 227  
 DB 168 MPALHTLNDHNLIDALPPGAFQOLGSLRLDLTSNRLATLAPDPLFSRGRDAEAPAPLV 227  
 QY 228 VLSFGNPLHCNCCELLMLRLARPDDLETCASPPGLAGRYFMAVPEGEFSCPEPLIARHT 287  
 DB 228 VLSFGNPLHCNCCELLMLRLARPDDLETCASPPGLAGRYFMAVPEGEFSCPEPLIARHT 287  
 QY 288 ORLWVLEGRATLRCRAIGDPAPTMHWVPPDRLVNSSRAAFPNGLTIEGYTGAGDAG 347  
 DB 288 ORLWVLEGRATLRCRAIGDPAPTMHWVPPDRLVNSSRAAFPNGLTIEGYTGAGDAG 347  
 QY 348 GYTCTATNPAGATATRAVELRVIALPHGNSAEGGRPGPSDIAASARTAEGBGTLESPP 407  
 DB 348 GYTCTATNPAGATATRAVELRVIALPHGNSAEGGRPGPSDIAASARTAEGBGTLESPP 407  
 QY 408 AVQTEVATSGVSWGGRPADPVMFQIYNSSEDETLIYRIYVASSHFLKHLVPG 467  
 DB 408 AVQTEVATSGVSWGGRPADPVMFQIYNSSEDETLIYRIYVASSHFLKHLVPG 467  
 QY 468 ADVYDCLIALSPAAPSDITATRLIGCAHFTLPAIPCHALQAVHGLGTLTVAAVGVV 527  
 DB 468 ADVYDCLIALSPAAPSDITATRLIGCAHFTLPAIPCHALQAVHGLGTLTVAAVGVV 527  
 QY 528 AALLVFTVALIVRGAGNGRLPLKLSHVQSQTNCGSPPTPKAHPRPSPRPORSCSLD 587  
 DB 528 AALLVFTVALIVRGAGNGRLPLKLSHVQSQTNCGSPPTPKAHPRPSPRPORSCSLD 587

```

Db      421  AALLVETVALVGRGANGRLPLKLSHVQSCNTGSGSPTRKXAPPRPPRQSCSLD 480
Qy      588  LGDAGCYGVARRLGAMARRSHSVHGGLLGAGCRGVGSAERLEESVY 635
Db      481  LGDAGCYGVARRLGAMARRSHSVHGGLLGAGCRGVGSAERLEESVY 528

RESULT 7
Q6PK41 ID Q6PK41 PRELIMINARY; PRT; 541 AA.
AC Q6PK41;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE LRFN4 protein (Fragment).
GN Name=LRFN4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McKernan P.U., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007718; AAH07718.1; -like.
DR InterPro; IPR006957; FW_III-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_cterm.
DR InterPro; IPR003591; LRR_cyp.
DR Pfam; PF00047; I9; 1.
DR Pfam; PF00560; LRR; 7.
DR PRINTS; PR00019; LRRICHRPT.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00369; LRR_Typ; 6.
DR PROSITE; PS50835; IG_LIKE; 1.
DR NON TER 541 541
SQ SEQUENCE 541 AA; 58167 MW; E01B96549DFE8E4F CRC64;
Query March 69.0%; Score 2302.5; DB 2; Length 541;
Best Local Similarity 90.9%; Pred. No. 2e-135;
Matches 450; Conservative 5; Mismatches 21; Indels 19; Gaps 3;
Qy 1 MAPPLILLILLASGAACPLPCVCQNLSESLSTLCAHRGLTFVPPVNDRTVELRLADNFI 60

```

```

Db      1  MAPPLILLILLASGAACPLPCVCQNLSESLSTLCAHRGLTFVPPVNDRTVELRLADNFI 60
Qy      61  QALGPPDFRNMTGLVDLTLSRNAITTRIGARAFGDLSELSLHDGNRLVELGTGSLRGVY 120
Db      61  QALGPPDFRNMTGLVDLTLSRNAITTRIGARAFGDLSELSLHDGNRLVELGTGSLRGVY 120
Qy      121  NIQHLITSGNQGRIRIPGAFDDPLESLEDLDSYNNLRQVPMAGIGAMPRLHTLNDHNL 180
Db      121  NIQHLITSGNQGRIRIPGAFDDPLESLEDLDSYNNLRQVPMAGIGAMPRLHTLNDHNL 180
Qy      181  IDALPPGAPAOIGQLSRDLTNSRLATLAPDPLFSSGRDAEAPAPLVISFSGNPLHCNC 240
Db      181  IDALPPGAPAOIGQLSRDLTNSRLATLAPDPLFSSGRDAEAPAPLVISFSGNPLHCNC 240
Qy      241  ELLIMRLARPDDLFCASPPGAGRYFMAVPEGEFSCEPPLIARHTQRLMTVEGORATL 300
Db      241  ELLIMRLARPDDLFCASPPGAGRYFMAVPEGEFSCEPPLIARHTQRLMTVEGORATL 300
Qy      301  RCRAIDGPAPVTMVGPPDDLGVNSSRARAFPNGTLEIGVTGAGDAGYTCIATNPAGEA 360
Db      301  RCRAIDGPAPVTMVGPPDDLGVNSSRARAFPNGTLEIGVTGAGDAGYTCIATNPAGEA 360
Qy      361  TAEVELRVIALPHGNSSAEGRPSPDIAASARTAAEGEGTLESEPAVQVTEVATSGL 420
Db      361  TAEVELRVIALPHGNSSAEGRPSPDIAASARTAAEGEGTLESEPAVQVTEVATSGL 420
Qy      421  VSMGPRPADPVMVFOIQYNNSSDEDTLYIVPASHHLLKTLVGCADVDTCLLALSP- 479
Db      421  VSMGPRPADPVMVFOIQ- RANGRRILWBSLPGFYSQLRLPCVLGG-----AGSPF 471
Qy      480  -----AAGPSD 485
Db      472  LGSGLQTRVAPAGPSN 486

RESULT 8
AAH07718 PRELIMINARY; PRT; 541 AA.
AC AAH07718;
DT 02-MAR-2004 (TEMBLrel. 27, Created)
DT 02-MAR-2004 (TEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TEMBLrel. 27, Last annotation update)
DE LRFN4 protein (Fragment).
GN LRFN4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McKernan P.U., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```

[2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RA Strausberg R.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, BC007718; AA07718.1; -  
 FT NON\_TER 541 541  
 SQ SEQUENCE 541 AA; 58167 MW; E01B98549DF6BE4F CRC64;  
 Query Match 69.0%; Score 2302.5; DB 2; Length 541;  
 Best Local Similarity 90.9%; Pred. No. 2e-135;  
 Matches 450; Conservative 5; Mismatches 21; Indels 19; Gaps 3;  
 QY 1 MAPPILLILLASGAACPLPCVCQNTSESLSTLCAHRLGFVPPNVDRTVELRLADNFI 60  
 DB 1 MAPPILLILLASGAACPLPCVCQNTSESLSTLCAHRLGFVPPNVDRTVELRLADNFI 60  
 QY 61 QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHDGNRLVELGTGSLRGPV 120  
 DB 61 QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHDGNRLVELGTGSLRGPV 120  
 QY 121 NIQHLLISGNQGRIPAGAFDDFLESLEDDLSYNNLRQVPMAGIGAMPALHTLNDHNL 180  
 DB 121 NIQHLLISGNQGRIPAGAFDDFLESLEDDLSYNNLRQVPMAGIGAMPALHTLNDHNL 180  
 QY 181 IDALPPGAPAOIGQLSRDLTNSRLATLAPDPLFSRGRDAESPAPVLVSFGNPLHCNC 240  
 DB 181 IDALPPGAPAOIGQLSRDLTNSRLATLAPDPLFSRGRDAESPAPVLVSFGNPLHCNC 240  
 QY 241 ELIMRLRLARPDLETCASPPGLAGRYFMAVPEGEFSCPEPLIARTQRLMWLEGQRATL 300  
 DB 241 ELIMRLRLARPDLETCASPPGLAGRYFMAVPEGEFSCPEPLIARTQRLMWLEGQRATL 300  
 QY 301 RCALGDPAPPTMHWGPDDRLVGNSSRARAFPNGTLEIGVTAGDAGYTCTATNPAGEA 360  
 DB 301 RCALGDPAPPTMHWGPDDRLVGNSSRARAFPNGTLEIGVTAGDAGYTCTATNPAGEA 360  
 QY 361 TRAVELRVIALPHGNSAEGGRPGPSDIAASARTAEGETSESPAVQTEVTATSGI 420  
 DB 361 TRAVELRVIALPHGNSAEGGRPGPSDIAASARTAEGETSESPAVQTEVTATSGI 420  
 QY 421 VSKGPRPADPVMFMQIQNNSSEDEFTLIRIVPASHHPLKHLPGADYDLCLALSP- 479  
 DB 421 VSKGPRPADPVMFMQIQNNSSEDEFTLIRIVPASHHPLKHLPGADYDLCLALSP- 471  
 QY 480 -----AAGPSD 485  
 DB 472 LSGGLQTRVYRAGPSN 486  
 RESULT 9  
 Q6IPL6 PRELIMINARY; PRT; 460 AA.  
 AC 06IPL6;  
 DT 05-JUL-2004 (TREMblrel. 27, Created)  
 DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)  
 DE 05-JUL-2004 (TREMblrel. 27, Last annotation update)  
 OS Hypothetical protein (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Urdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McWay P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Rane J., Helton B., Kettelman W., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Gichwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalins D.E., Schermer A., Schein J.E.,  
 RA Jones S.J., Maira M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RA Strausberg R.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, BC071866; AA071866.1; -  
 DR InterPro: IPR003599; IG.  
 DR InterPro: IPR007110; IG-like.  
 DR InterPro: IPR003598; IG\_c2.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR000483; LRR\_Cterm.  
 DR InterPro: IPR003591; LRR\_Typ.  
 DR Pfam: PF00047; IG\_1.  
 DR Pfam: PF00560; LRR\_7.  
 DR PRINTS, PR00019; LEURICRPT.  
 DR SMART, SM00409; IG\_1.  
 DR SMART, SM00408; IG\_c2; 1.  
 DR SMART, SM00082; LRRCT; 1.  
 DR SMART, SM00369; LRR\_Typ; 6.  
 DR PROSITE, PS00835; IG\_Like; 1.  
 DR Hypothetical protein.  
 FT NON\_TER 460 460  
 SQ SEQUENCE 460 AA; 49273 MW; FBD686F0AD360040 CRC64;  
 Query Match 60.3%; Score 2013; DB 2; Length 460;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-117;  
 Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAPPILLILLASGAACPLPCVCQNTSESLSTLCAHRLGFVPPNVDRTVELRLADNFI 60  
 DB 1 MAPPILLILLASGAACPLPCVCQNTSESLSTLCAHRLGFVPPNVDRTVELRLADNFI 60  
 QY 61 QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHDGNRLVELGTGSLRGPV 120  
 DB 61 QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHDGNRLVELGTGSLRGPV 120  
 QY 121 NIQHLLISGNQGRIPAGAFDDFLESLEDDLSYNNLRQVPMAGIGAMPALHTLNDHNL 180  
 DB 121 NIQHLLISGNQGRIPAGAFDDFLESLEDDLSYNNLRQVPMAGIGAMPALHTLNDHNL 180  
 QY 181 IDALPPGAPAOIGQLSRDLTNSRLATLAPDPLFSRGRDAESPAPVLVSFGNPLHCNC 240  
 DB 181 IDALPPGAPAOIGQLSRDLTNSRLATLAPDPLFSRGRDAESPAPVLVSFGNPLHCNC 240  
 QY 241 ELIMRLRLARPDLETCASPPGLAGRYFMAVPEGEFSCPEPLIARTQRLMWLEGQRATL 300  
 DB 241 ELIMRLRLARPDLETCASPPGLAGRYFMAVPEGEFSCPEPLIARTQRLMWLEGQRATL 300  
 QY 301 RCALGDPAPPTMHWGPDDRLVGNSSRARAFPNGTLEIGVTAGDAGYTCTATNPAGEA 360  
 DB 301 RCALGDPAPPTMHWGPDDRLVGNSSRARAFPNGTLEIGVTAGDAGYTCTATNPAGEA 360  
 QY 361 TRAVELRVIALPHGNSAEGGR 383  
 DB 361 TRAVELRVIALPHGNSAEGGR 383  
 RESULT 10  
 AA071866 PRELIMINARY; PRT; 460 AA.  
 ID AA071866

AAH71866;  
 DT 01-JUN-2004 (TReMBLrel. 27, Created)  
 DT 01-JUN-2004 (TReMBLrel. 27, Last sequence update)  
 DT 01-JUN-2004 (TReMBLrel. 27, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RA Strausberg R.;  
 RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; BC071866; AAH71866.1;  
 KW Hypothetical protein.  
 FT NON TER 460 460  
 SQ SEQUENCE 460 AA; 49273 MW; FBD686F0AD360040 CRC64;  
 Query Match 60.3%; Score 2013; DB 2; Length 460;  
 Best Local Similarity 100.0%; Pred. No. 1,9e-117;  
 Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAPPILLILLASGAACPLPCVCONLSSELSSTICAHGGLFVPPNVDRTVELRLADNFI 60  
 DB 1 MAPPILLILLASGAACPLPCVCONLSSELSSTICAHGGLFVPPNVDRTVELRLADNFI 60  
 QY 61 QALGPPFRFMTGVLDTLTSRNATIRGARAFGLSRLSHLDGNLVGLGSLRGPV 120  
 DB 61 QALGPPFRFMTGVLDTLTSRNATIRGARAFGLSRLSHLDGNLVGLGSLRGPV 120  
 QY 121 NLQHLISGNQGLRIAGAFDDFLSLEDLSTYNNLRQVWAGIGAMPALHTLNDNHL 180  
 DB 121 NLQHLISGNQGLRIAGAFDDFLSLEDLSTYNNLRQVWAGIGAMPALHTLNDNHL 180  
 QY 181 IDALPPGAFQOLGSLDLSNRLATLADPPLFSRGRDABASAPLVLSFGNPLHNC 240  
 DB 181 IDALPPGAFQOLGSLDLSNRLATLADPPLFSRGRDABASAPLVLSFGNPLHNC 240  
 QY 241 ELMLRLRLAPDDLETASPPGLAGRFMAVPSGESECEPPLTARHQRMLVLEGQATL 300  
 DB 241 ELMLRLRLAPDDLETASPPGLAGRFMAVPSGESECEPPLTARHQRMLVLEGQATL 300  
 QY 301 RCALGDPAPTMWVGDDRLVGNSSRRARAFPNGLTLEIGVTGADAGYTCIATNPAGEA 360  
 DB 301 RCALGDPAPTMWVGDDRLVGNSSRRARAFPNGLTLEIGVTGADAGYTCIATNPAGEA 360  
 QY 361 TARVELRVLALPHGNGSSAEGGR 383

DB 361 TARVELRVLALPHGNGSSAEGGR 383  
 RESULT 11  
 ID 0713C2 PRELIMINARY; PRT; 329 AA.  
 AC 0713C2;  
 DT 05-JUL-2004 (TReMBLrel. 27, Created)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
 DE LRPN4 protein (Fragment).  
 OS Homo sapiens (Human).  
 GN Name=LRPN4;  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon;  
 RA Strausberg R.;  
 RA Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; BC014040; AAH14040.2;  
 DR InterPro; IPR003961; FN\_III.  
 DR InterPro; IPR008957; FN\_III-like.  
 DR Pfam; PF00041; fn3; 1.  
 DR SMART; SM00060; FN3; 1.  
 DR PROSITE; PSS0853; FN3; 1.  
 DR PROSITE; PSS0835; IG-LIKE; 1.  
 FT NON TER 1 1  
 SQ SEQUENCE 329 AA; 33674 MW; 0E78B9038E67DB8 CRC64;  
 Query Match 51.9%; Score 1733; DB 2; Length 329;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-100;  
 Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 307 DPAPTMWVGDDRLVGNSSRRARAFPNGLTLEIGVTGADAGYTCIATNPAGEATARVEL 366  
 DB 1 DPAPTMWVGDDRLVGNSSRRARAFPNGLTLEIGVTGADAGYTCIATNPAGEATARVEL 366  
 QY 367 RVIALPHGNGSSAEGGRPGSDIAAGARTABEGTLESFPAVQVBTATSGLVSKPG 426  
 DB 61 RVIALPHGNGSSAEGGRPGSDIAAGARTABEGTLESFPAVQVBTATSGLVSKPG 426  
 QY 427 RPADPVWMPQIOYNSSEDFLIRIVPASSHHFLKHLVGAUYDCLALSPAAGPSDL 486  
 DB 121 RPADPVWMPQIOYNSSEDFLIRIVPASSHHFLKHLVGAUYDCLALSPAAGPSDL 486  
 QY 487 TARLLGCAHFTSLPASPLCHALQAHVLGTLTVAVGVVLVAALLVFTVALLVGRGAGN 546

```

Db      181 TATRLGCAHFSTLPASPLCHALQAHVGGTLTVAVGVVVAALVFTVALLVGRGAGN 240
Qy      547 GRPLKLSHVOSQTNQSPSPPTPKAHPRSPRPSRSCSLDLDAGCGYARRLGAGMAR 606
Db      241 GRPLKLSHVOSQTNQSPSPPTPKAHPRSPRPSRPSRSCSLDLDAGCGYARRLGAGMAR 300
Qy      607 RSHSVHGILLGAGCGVGGSAERLEESV 635
Db      301 RSHSVHGILLGAGCGVGGSAERLEESV 329

RESULT 12
AAH14040
ID AAH14040 PRELIMINARY; PRT; 329 AA.
AC AAH14040;
DT 02-MAR-2004 (TEMBLrel. 27, Created)
DT 02-MAR-2004 (TEMBLrel. 27, Last sequence update)
DE 02-MAR-2004 (TEMBLrel. 27, Last annotation update)
DE LRFN4 protein (Fragment).
GN LRFN4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kerteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014040; AAH14040.2; -.
FT NON TER
SQ SEQUENCE 329 AA; 33674 MW; 0E78B9038BE67DB8 CRC64;

Query Match 51.9%; Score 1733; DB 2; Length 329;
Best Local Similarity 100.0%; Pred. No. 3.8e-100; Indels 0; Gaps 0;
Matches 329; Conservative 0; Mismatches 0;

Db      307 DPAPTMHWGPDRLVGNSSRRARAFPNGLTLEIGVTGAGDAGGYTCIATNPAGEATARVEL 366
Qy      367 RYALPHGNSASBEGRPSPDIASARTAAEGTLESEPVQVTEVTATISGLVSGPG 426
Db      61 RYALPHGNSASBEGRPSPDIASARTAAEGTLESEPVQVTEVTATISGLVSGPG 120
Qy      427 RPADPVMWFQIQYNSSEDETLIRIVPASSHHFLKHLVPGADYDLCILASPAGSPDL 486
Db      121 RPADPVMWFQIQYNSSEDETLIRIVPASSHHFLKHLVPGADYDLCILASPAGSPDL 180

```

```

Qy      487 TATRLGCAHFSTLPASPLCHALQAHVGGTLTVAVGVVVAALVFTVALLVGRGAGN 546
Db      181 TATRLGCAHFSTLPASPLCHALQAHVGGTLTVAVGVVVAALVFTVALLVGRGAGN 240
Qy      547 GRPLKLSHVOSQTNQSPSPPTPKAHPRSPRPSRSCSLDLDAGCGYARRLGAGMAR 606
Db      241 GRPLKLSHVOSQTNQSPSPPTPKAHPRSPRPSRPSRSCSLDLDAGCGYARRLGAGMAR 300
Qy      607 RSHSVHGILLGAGCGVGGSAERLEESV 635
Db      301 RSHSVHGILLGAGCGVGGSAERLEESV 329

RESULT 13
Q9BWJ0
ID Q9BWJ0 PRELIMINARY; PRT; 327 AA.
AC Q9BWJ0;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DE 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE LRFN4 protein (Fragment).
GN LRFN4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kerteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC000207; AAH00207.2; -.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00041; fn3; 1.
DR SMART; SMO0060; FN3; 1.
DR PROSITE; PS0053; FN3; 1.
FT NON TER
SQ SEQUENCE 327 AA; 33461 MW; A5582938BEF197214 CRC64;

Query Match 51.6%; Score 1720; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 2.4e-99; Indels 0; Gaps 0;
Matches 327; Conservative 0; Mismatches 0;

Db      309 APTMHWGPDRLVGNSSRRARAFPNGLTLEIGVTGAGDAGGYTCIATNPAGEATARVEL 368
Qy      1 APTMHWGPDRLVGNSSRRARAFPNGLTLEIGVTGAGDAGGYTCIATNPAGEATARVEL 60

```



```

QY 369 LALPHGNSAAGRGPGSDIAASARTAAEGGTLESBAVQVTEVTATSGVSWGPGRP 428
DB 61 LALPHGNSAAGRGPGSDIAASARTAAEGGTLESBAVQVTEVTATSGVSWGPGRP 120
QY 429 ADPVWMEFOIQVNSSEDETLIYRIVPASSHHFLKHLVPGADYDLCILALSPAAGSDLTAA 488
DB 121 ADPVWMEFOIQVNSSEDETLIYRIVPASSHHFLKHLVPGADYDLCILALSPAAGSDLTAA 180
QY 489 TRLLGCAHFSTLPASPLCHALQAHVLGGLTVAVGVVAALLVFTVALLVGRGANGR 548
DB 181 TRLLGCAHFSTLPASPLCHALQAHVLGGLTVAVGVVAALLVFTVALLVGRGANGR 240
QY 549 LPLKLSHVOSQTNNGGSPPTPKAHPSPSPRRPQSCSLDGDAGCYGVARLGGAMARRS 608
DB 241 LPLKLSHVOSQTNNGGSPPTPKAHPSPSPRRPQSCSLDGDAGCYGVARLGGAMARRS 300
QY 609 HSVHGGILGAGCRGVGSAERLEESVY 635
DB 301 HSVHGGILGAGCRGVGSAERLEESVY 327

```

## RESULT 14

```

ID AAH00207 PRELIMINARY; PRT; 327 AA.
AC AAH00207.

```

```

DT 02-MAR-2004 (TEMBLrel. 27, Created)
DT 02-MAR-2004 (TEMBLrel. 27, Last sequence update)
DE LRFN4 protein (Fragment).
GN LRFN4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhany S.J.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whaley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzanski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC000207; AAH00207.2; -.
FT NON TER 1
SQ SEQUENCE 327 AA; 33461 MW; A5582938BF197214 CRC64;

```

```

Query Match 51.6%; Score 1720; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. NO. 2.4e-99;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
309, APTMHWGPDRLVGNSSRAAFPNGLTIGVTGAGDAGGYTCIATNPAGATARVELRV 368

```

```

DB 1 APTMHWGPDRLVGNSSRAAFPNGLTIGVTGAGDAGGYTCIATNPAGATARVELRV 60
QY 369 LALPHGNSAAGRGPGSDIAASARTAAEGGTLESBAVQVTEVTATSGVSWGPGRP 428
DB 61 LALPHGNSAAGRGPGSDIAASARTAAEGGTLESBAVQVTEVTATSGVSWGPGRP 120
QY 429 ADPVWMEFOIQVNSSEDETLIYRIVPASSHHFLKHLVPGADYDLCILALSPAAGSDLTAA 488
DB 121 ADPVWMEFOIQVNSSEDETLIYRIVPASSHHFLKHLVPGADYDLCILALSPAAGSDLTAA 180
QY 489 TRLLGCAHFSTLPASPLCHALQAHVLGGLTVAVGVVAALLVFTVALLVGRGANGR 548
DB 181 TRLLGCAHFSTLPASPLCHALQAHVLGGLTVAVGVVAALLVFTVALLVGRGANGR 240
QY 549 LPLKLSHVOSQTNNGGSPPTPKAHPSPSPRRPQSCSLDGDAGCYGVARLGGAMARRS 608
DB 241 LPLKLSHVOSQTNNGGSPPTPKAHPSPSPRRPQSCSLDGDAGCYGVARLGGAMARRS 300
QY 609 HSVHGGILGAGCRGVGSAERLEESVY 635
DB 301 HSVHGGILGAGCRGVGSAERLEESVY 327

```

```

ID OBN644 PRELIMINARY; PRT; 324 AA.
AC OBN644.

```

```

DT 01-OCT-2002 (TEMBLrel. 22, Created)
DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DE LRFN4 protein.
GN LRFN4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhany S.J.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whaley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzanski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027475; AAH27475.1; -.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00041; fn3; 1.
DR SMART; SM00060; FN3; 1.
DR PROSITE; PS50853; FN3; 1.

```

DR PROSITE; PS50835; IG\_LIKE; 1.  
SQ SEQUENCE 324 AA; 33164 MW; 6DC6E0871E227E2A CRC64;

Query Match 51.0%; Score 1700; DB 2; Length 324;  
Best Local Similarity 99.7%; Pred. No. 4.3e-98;

Matches 323; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 312 MHWVGPDRLVGNSSRAAFPNGLTEIGVTGADGAGYTCIATNPAGEATARVELRVLAI 371
    |||
Db 1 MHWVGPDRLVGNSSRAAFPNGLTEIGATGADGAGYTCIATNPAGEATARVELRVLAI 60
    |||
QY 372 PHGNSSAEGGRPGPSDIAASARTAAAGSGTLESEPAVQYEVYATSGLVSWGGRPADP 431
    |||
Db 61 PHGNSSAEGGRPGPSDIAASARTAAEGEGTLESEPAVQYEVYATSGLVSWGGRPADP 120
    |||
QY 432 VMWFQIYNSSSEDETLIYRIIVPASSHHFLKHLVPGADYDLCILALSPAAGPSDLTATRL 491
    |||
Db 121 VMWFQIYNSSSEDETLIYRIIVPASSHHFLKHLVPGADYDLCILALSPAAGPSDLTATRL 180
    |||
QY 492 LGCARFSTLPASPLCHALQAHVLGTLTVAVGGLVAALLVFTVALLVGRGAGNGRLPL 551
    |||
Db 181 LGCARFSTLPASPLCHALQAHVLGTLTVAVGGLVAALLVFTVALLVGRGAGNGRLPL 240
    |||
QY 552 KLSHVQSQTNGGSPPTPKAHPSPSPPPQPSCSLDLSDAGCYGYARRLGAWARRSHSV 611
    |||
Db 241 KLSHVQSQTNGGSPPTPKAHPSPSPPPQPSCSLDLSDAGCYGYARRLGAWARRSHSV 300
    |||
QY 612 HGGILGAGCGRGVGSARLEESYV 635
    |||
Db 301 HGGILGAGCGRGVGSARLEESYV 324
    |||
```

Search completed: November 17, 2004, 22:40:08  
Job time : 201 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: November 17, 2004, 22:30:58 / Search time 73 Seconds  
(without alignments) 3120.453 Million cell updates/sec

Title: US-10-071-879-10

Perfect score: 3336  
Sequence: 1 MAPPLDLLLILASGAAACPLP.....LGAGCRGVGSAERLEESVY 635

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

1: A\_Geneseq\_23Sep04:\*  
2: geneseqp1980s:\*  
3: geneseqp1990s:\*  
4: geneseqp2000s:\*  
5: geneseqp2001s:\*  
6: geneseqp2002s:\*  
7: geneseqp2003as:\*  
8: geneseqp2003bs:\*  
9: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3331	99.9	635	6	AAO26256 MDT rela
2	3330	99.8	635	5	AAE23980 Human LP2
3	3219	96.5	778	5	ABP70144 Human NOV
4	3218.5	96.5	618	7	AD121104
5	3026.5	90.7	647	5	ABP70142 Human NOV
6	2970	89.0	565	7	ADP08361 Novel pro
7	2773.5	83.1	566	5	ABP70143 Human NOV
8	2773	83.1	551	5	AAE17484 Human leu
9	2568.5	77.0	526	4	ABG04827 Novel hum
10	2475	74.2	468	4	AA870072 Human sec
11	2475	74.2	468	5	ABG55511 Human alb
12	2475	74.2	468	8	AD178778 Human hno
13	1662.5	49.8	785	3	AA812448 Human bira
14	1662.5	49.8	789	3	AA809968 Human MP5
15	1662.5	49.8	789	7	ADP69106 Human MP5
16	1655.5	49.6	789	4	AA839059 Human pol
17	1655.5	49.6	789	4	AA839059 Human pol
18	1564.5	46.9	636	4	AAU28270 Novel hum
19	1557	46.7	628	4	AA655805 Human leu
20	1556	46.6	628	4	AA655805 Human leu
21	1556	46.6	628	4	AA655805 Human leu
22	1556	46.6	628	4	AA655805 Human leu
23	1556	46.6	628	4	AA655805 Human leu
24	1556	46.6	628	4	AA655805 Human leu
25	1556	46.6	628	4	AA655805 Human leu

26	1546.5	46.4	627	5	ABG34079	ABG34079 Human pro
27	1546.5	46.4	627	6	ADA01368	ADA01368 Human pro
28	1546.5	46.4	627	6	ADA43797	ADA43797 Human sec
29	1546.5	46.4	627	6	ADA43565	ADA43565 Human sec
30	1546.5	46.4	627	6	ADA01240	ADA01240 Human pro
31	1546.5	46.4	627	7	ADA01124	ADA01124 Human sec
32	1546.5	46.4	627	7	ADA43681	ADA43681 Human pro
33	1546.5	46.4	627	7	ADA06943	ADA06943 Human pro
34	1546.5	46.4	627	7	ADA08431	ADA08431 Human pro
35	1546.5	46.4	627	7	AD89724	AD89724 Human pro
36	1546.5	46.4	627	7	AD887007	AD887007 Human pro
37	1546.5	46.4	627	7	AD866162	AD866162 Human sec
38	1546.5	46.4	627	7	AD899840	AD899840 Human pro
39	1546.5	46.4	627	7	AD899495	AD899495 Novel hum
40	1546.5	46.4	627	7	AD866046	AD866046 Human sec
41	1546.5	46.4	627	7	AD823444	AD823444 Human tra
42	1546.5	46.4	627	7	AD826137	AD826137 Human pro
43	1546.5	46.4	627	7	AD804964	AD804964 Human pro
44	1546.5	46.4	627	7	AD811270	AD811270 Human pro
45	1546.5	46.4	627	7	AD888201	AD888201 Human pro

## ALIGNMENTS

RESULT 1	AAO26256	standard, protein, 635 AA.
ID	AAO26256	
XX	AAO26256	
AC	AAO26256	
XX	AAO26256	
DT	10-APR-2003	(first entry)
XX	10-APR-2003	
DE	MDT related human protein SEQ ID No 34.	
XX	MDT related human protein SEQ ID No 34.	
XX	Cystostatic; antiatherosclerotic; osteopathic; antiarteriosclerotic;	
KW	hepatostatic; antipapillary; antiallergic; antianemic; antiaesthetic;	
KW	antithyroid; antinflammatory; antihelminthic; antidiabetic; nephrotoxic;	
KW	ophthalmologic; immunosuppressive; dermatological; antifungal;	
KW	antihemorrhagic; antirheumatic; antidiabetic; antitumor;	
KW	antiparasitic; protozoacide; tranquilizer; antihiv;	
KW	neurotropic; neuroprotective; anticonvulsant; cerebroprotective;	
KW	neuroleptic; molecules for disease detection and treatment; MDT;	
KW	immunogen; cancer; actinic keratosis; arteriosclerosis; atherosclerosis;	
KW	burstitis; cirrhosis; hepatitis; psoriasis; AIDS; rheumatoid arthritis;	
KW	adult respiratory distress syndrome; Addison's disease; allergy; anaemia;	
KW	asthma; osteoporosis; autoimmune; haemolytic anaemia; scleroderma;	
KW	autoimmune thyroiditis; Crohn's disease; atopic dermatitis;	
KW	diabetes mellitus; Graves' disease; glomerulonephritis;	
KW	systemic lupus erythematosus; systemic sclerosis; ulcerative colitis;	
KW	haemodialysis; uveitis; trauma; Alzheimer's; Pick disease;	
KW	Parkinson disease; amyotrophic lateral sclerosis; epilepsy; stroke;	
KW	Huntington's disease; multiple sclerosis; dementia;	
KW	extrapyramidal disorder; motor neuron disorder; central nervous system;	
KW	neuromuscular disorder; metabolic; endocrine; toxic myopathy;	
XX	periodic paralysis; mental disorder; human.	
OS	Homo sapiens.	
XX	WO200296951-A1.	
PN	05-DEC-2002.	
XX	24-MAY-2002; 2002WO-US016676.	
PF	25-MAY-2001; 2001US-0293723P.	
XX	01-JUN-2001; 2001US-0295257P.	
PR	08-JUN-2001; 2001US-0297220P.	
PR	21-JUN-2001; 2001US-0300526P.	
PR	29-JUN-2001; 2001US-0301874P.	
PR	22-FEB-2002; 2002US-0359413P.	
XX	(INCY-) INCYTE GENOMICS INC.	
PA		

XX Tang TY, Yue H, Baughn MR, Duggan BM, Warren BA, Bandman O;  
 PI Richardson TW, Burford N, Santhamala B, Becha SD, Yao MG, Yang J;  
 PI Tiran UK, Hafellia AUZ, Griffin JA, Swannakr A, Elliott VS;  
 PI Recton SA, Khan FA, Lee EA, Yue H, Lu DM, Walla NK, Thangavelu K;  
 PI Ariyazu CS, Xu Y, Ison CH, Huang J, Ding L, Honchell CD;  
 PI Borowsky ML, Emerling BM, Peterson DP, Lu Y, Ramkumar J, Mason PM;  
 PI Zeharjadian Y, Azimzai Y, Stuve LL, Kamigaki LL, Barroso I, Lee S;  
 PI Kable AE;  
 XX  
 DR MPI: 2003-140448/13.  
 DR N-PSDB: AAK99609.  
 XX  
 PT Novel molecules for disease detection and treatment and polynucleotide  
 PT encoding them useful for diagnosing, preventing or treating cell  
 PT proliferative, autoimmune/inflammatory, neurological and developmental  
 PT disorders.  
 XX  
 S8 Claim 89, Page 218-220; 260pp; English.

The invention relates to an isolated polypeptide chosen from molecules for disease detection and treatment (MDMT), comprising a one of 39 114-1250 residue amino acid sequences, given in the specification, or a biologically active or immunogenic fragment of the isolated polypeptide. The isolated polypeptide is useful for screening a compound for effectiveness as an agonist or antagonist of the isolated polypeptide. The isolated polypeptide is also useful as an immunogen for preparing polyclonal or monoclonal antibodies by hybridoma technology. The isolated polypeptide and its encoding polynucleotide are useful for diagnosis, treatment and prevention of cancer; actinic keratosis, arteriosclerosis, atherosclerosis, burns, cirrhosis, hepatitis, psoriasis, AIDS, adult respiratory distress syndrome, Addison's disease, allergies, anaemia, asthma, atherosclerosis, osteoporosis, autoimmune haemolytic anaemia, autoimmune thyroiditis, Crohn's disease, atopic dermatitis, diabetes mellitus, Graves' disease, glomerulonephritis, rheumatoid arthritis, scleroderma, systemic lupus erythematosus, systemic sclerosis, ulcerative colitis, haemodialysis, uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic infections, trauma, Alzheimer's and Pick disease, Parkinson disease, amyotrophic lateral sclerosis, epilepsy, stroke, Huntington's disease, multiple sclerosis, dementia and other extrapyramidal disorder, motor neuron disorder, and other developmental disorders of the central nervous system, neuromuscular disorders, metabolic, endocrine and toxic myopathies, periodic paralyses, mental disorders including mood, anxiety and schizophrenic disorders, anaemia, renal tubular acidosis, epilepsy, hypothyroidism, glaucoma, sensorineural hearing loss and cataract. This sequence represents a human MDMT protein relating to the invention

SQ Sequence 635 AA;

Query Match	99.9%	Score 3331	DB 6	length 635
Best Local Similarity	99.8%	Pred. No. 2e-250		
Matches 634; Conservative	1	Mismatches	0	Indels 0; Gaps 0;

QY	1	MAPPELLILLIAGAAACAPCYCONISELSITCARGILLFPPANDRSTVELRLADNFI	60
Db	1	MAPPELLILLIAGAAACAPCYCONISELSITCARGILLFPPANDRSTVELRLADNFI	60
QY	61	QALGPPDFRNTGTVDLTISRNAITRIGARFGLDESLSLHIDGNRLVELGTGSLRGVY	120
Db	61	QALGPPDFRSMTGVDLTLISRNAITRIGARFGLDESLSLHIDGNRLVELGTGSLRGVY	120
QY	121	NIQGHILISGNQIGRIAPGAFDPFLSELDLDSYNNLRQVPMAGIGAMPALHTLINDHN	180
Db	121	NIQGHILISGNQIGRIAPGAFDPFLSELDLDSYNNLRQVPMAGIGAMPALHTLINDHN	180
QY	181	IDALPPGAFAQLGQSLRDLTNSNRATTLPDPLFSHGRDAEASPADLVTSFSGNPLHCNC	240
Db	181	IDALPPGAFAQLGQSLRDLTNSNRATTLPDPLFSHGRDAEASPADLVTSFSGNPLHCNC	240
QY	241	ELIIMRRLARPPDLTCSAPGAGGYPMVAVBGESSCPPLIARTQELMTVEGGRTLT	300
Db	241	ELIIMRRLARPPDLTCSAPGAGGYPMVAVBGESSCPPLIARTQELMTVEGGRTLT	300

QY	301	RCRLGDPAIPMTMWWGPPDRLIVGNSSRRAPFNGTLEICVTASGAGAGYTCIATNPAGA	360
Db	301	RCRLGDPAIPMTMWWGPPDRLIVGNSSRRAPFNGTLEICVTASGAGAGYTCIATNPAGA	360
QY	361	TARVELRYALPHGNSSSAEGGRPGPSDIASARTAAEGGTLSESPAVOYTEVTATSGI	420
Db	361	TARVELRYALPHGNSSSAEGGRPGPSDIASARTAAEGGTLSESPAVOYTEVTATSGI	420
QY	421	VSMGGRPADPVMWFOIQYNSSSEDETLIRIYIPASSHHFLKHLVPGADYDLCCLALSPA	480
Db	421	VSMGGRPADPVMWFOIQYNSSSEDETLIRIYIPASSHHFLKHLVPGADYDLCCLALSPA	480
QY	481	AGPSLTLTRLILGCHHFSTLPASPLCLHALQAVILGTLTVANGVYVALILVFTVALVR	540
Db	481	AGPSLTLTRLILGCHHFSTLPASPLCLHALQAVILGTLTVANGVYVALILVFTVALVR	540
QY	541	GRGAGNGRLPLKLISVQSOTNGSPPTPKAHPRRSPPRQSCSLDLGDAGCYGARRL	600
Db	541	GRGAGNGRLPLKLISVQSOTNGSPPTPKAHPRRSPPRQSCSLDLGDAGCYGARRL	600
QY	601	GGAMARRSHSVHGLIGAGCRGVGSAERLIESVV	635
Db	601	GGAMARRSHSVHGLIGAGCRGVGSAERLIESVV	635

RESULT 2	
AAE23980	
ID	AAE23980 standard; protein; 635 AA
XX	
AC	AAE23980;
XX	
DT	23-SEP-2002 (first entry)

Human LP220 secreted protein.

KW Human; secreted protein; atherosclerosis; Alzheimer's disease; LP220;  
 KW diabetic retinopathy; severe combined immunodeficiency; pancreatitis;  
 KW rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;  
 KW reperfusion injury; arteriosclerosis; wound healing; transgenic animal;  
 KW gene therapy; neoplasm; transgenic; psoriasis; ischaemia; carcinoma;  
 KW chromosome 11q13.  
 XX  
 XX  
 OS Homo sapiens.

*Homo sapiens.*

FH	Key	Location/Qualifiers
FT	Peptide	1..16
FT		/label= Signal_peptide
FT	Protein	17..635
FT		/note= "Mature human Lp220 secreted protein"

PN WO200226801-A2.

PD 04-APR-2002.

PF 14-SEP-2001; 2001WO-US026026.

PR 28-SEP-2000; 2000US-023

PA (ELIL ) LILLY & CO ELI.

PI Su EW, Wang H;  
xx

DR WPI; 2002-471259/50.

XX  
XX

PT treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis, diabetes mellitus, arterio-sclerosis, ischaemia or reperfusion injury

PS Claim 8; Page 127-129; 145pp; English.

CC The invention relates to human secreted polypeptides designated LP095,

CC LP191, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic acid  
CC molecules encoding such polypeptides. Novel secreted proteins of the  
CC invention are used for treating diseases such as atherosclerosis,  
CC Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis,  
CC arteriosclerosis, rheumatoid arthritis, colorectal adenoma, severe  
CC combined immunodeficiency, ischaemia, carcinoma, haemolytic anaemia,  
CC reperfusion injury, neoplasms and cancer especially liver cancer. They  
CC are also used for wound healing. Polynucleotides of the invention can be  
CC used to generate transgenic animals or knock out animals, which in turn,  
CC are useful in the development and screening of therapeutically useful  
CC reagents for use in the treatment of diseases associated with LP  
CC polypeptide associated activity. They are also used in gene therapy. The  
CC present sequence is human LP220 secreted protein. LP220 gene is located  
CC on chromosome 11q13  
XX  
XX  
SQ Sequence 635 AA;  
Query Match 99.8%; Score 3330; DB 5; Length 635;  
Best Local Similarity 99.8%; Pred. No. 2.3e-250;  
Matches 634; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MAPPLLILLASGAACPVCVCONSESLSTICAHGILFVPPNDRVVELRLDNFI 60  
Db 1 MAPPLLILLASGAACPVCVCONSESLSTICAHGILFVPPNDRVVELRLDNFI 60  
QY 61 QALGPDFFRMGTGLVDLTLSRNAITRIGARAFGDLSESLHLDGNRLVELGTGSLRPV 120  
Db 61 QALGPDFFRMGTGLVDLTLSRNAITRIGARAFGDLSESLHLDGNRLVELGTGSLRPV 120  
QY 121 NLQHLILSGNQLRIAGADDDLESDLDLSYNNIRQVPAGIGAMPALHTLNDHNI 180  
Db 121 NLQHLILSGNQLRIAGADDDLESDLDLSYNNIRQVPAGIGAMPALHTLNDHNI 180  
QY 181 IDALPGAFAGQLRLDLSNRILATLADPLFSRGRDSEAPALVIFSNGNPLHCNC 240  
Db 181 IDALPGAFAGQLRLDLSNRILATLADPLFSRGRDSEAPALVIFSNGNPLHCNC 240  
QY 241 ELLMLRRLARPDDLETCAAPGLAGRYFWAVPEGEFSCEPPLIARHTQRLMVEGQRATL 300  
Db 241 ELLMLRRLARPDDLETCAAPGLAGRYFWAVPEGEFSCEPPLIARHTQRLMVEGQRATL 300  
QY 301 RCALADDPATMMVGPDDRLVGNSSRARAFENGTLEIGVTGADGAGYTCIATNPAGEA 360  
Db 301 RCALADDPATMMVGPDDRLVGNSSRARAFENGTLEIGVTGADGAGYTCIATNPAGEA 360  
QY 361 TARVELRVIALPHGNSAAGGRPGPSDILAAARTAAEGGTESEPAVQVTEVTATSGI 420  
Db 361 TARVELRVIALPHGNSAAGGRPGPSDILAAARTAAEGGTESEPAVQVTEVTATSGI 420  
QY 421 VSWGPRPADPVMMFOIYVNSSEDETLIRIVPASSHFLKILVPGADVDLCILALSPA 480  
Db 421 VSWGPRPADPVMMFOIYVNSSEDETLIRIVPASSHFLKILVPGADVDLCILALSPA 480  
QY 481 AGPSDITARRLLGCAFFSTLPASPLCHAIQAHVLTGTLVAVGVVIAALVETVALVR 540  
Db 481 AGPSDITARRLLGCAFFSTLPASPLCHAIQAHVLTGTLVAVGVVIAALVETVALVR 540  
QY 541 GRGAGNGRLPLKLSHVOSQTNNGSPPTPKAHPRSPRPORCSLDIAGCYGARLT 600  
Db 541 GRGAGNGRLPLKLSHVOSQTNNGSPPTPKAHPRSPRPORCSLDIAGCYGARLT 600  
QY 601 GGAMARRSHSVHGILLAGCGRGVGSAAERLESSEVV 635  
Db 601 GGAMARRSHSVHGILLAGCGRGVGSAAERLESSEVV 635  
RESULT 3  
ABP70144  
ID ABP70144 standard; protein; 778 AA.  
XX  
AC ABP70144;  
XX  
DT 27-JAN-2003 (first entry)

XX  
DE Human NOV44C.  
XX  
KW Human; anti-HIV; cytostatic; antidiabetic; antiaesthetic; cachexia; AIDS;  
KW antiinflammatory; cardiac; haemostatic; neuroprotective; anorectic;  
KW neurotropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;  
KW antileukemic; cerebroprotective; gene therapy; NOVX; NOV; fertility;  
KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;  
KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; cardiovascular disorder;  
KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;  
KW metabolic syndrome X; wasting disorder; cell differentiation;  
KW cell proliferation; haematopoiesis; wound healing; angiogenesis.  
XX  
OS Homo sapiens.  
XX  
PN WO200272771-A2.  
XX  
PD 19-SEP-2002.  
XX  
PF 08-MAR-2002; 2002WC-US007288.  
XX  
PR 08-MAR-2001; 2001US-0274101P.  
XX  
PR 08-MAR-2001; 2001US-0274194P.  
PR 08-MAR-2001; 2001US-0274281P.  
PR 08-MAR-2001; 2001US-0274322P.  
PR 09-MAR-2001; 2001US-0274849P.  
PR 12-MAR-2001; 2001US-0275235P.  
PR 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
PR 13-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
PR 16-MAR-2001; 2001US-0276776P.  
PR 20-MAR-2001; 2001US-0276894P.  
PR 20-MAR-2001; 2001US-0277239P.  
PR 20-MAR-2001; 2001US-0277321P.  
PR 20-MAR-2001; 2001US-0277327P.  
PR 20-MAR-2001; 2001US-0277338P.  
PR 21-MAR-2001; 2001US-0277791P.  
PR 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
PR 26-MAR-2001; 2001US-0278894P.  
PR 27-MAR-2001; 2001US-0278936P.  
PR 27-MAR-2001; 2001US-0279036P.  
PR 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0279895P.  
PR 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280802P.  
PR 02-APR-2001; 2001US-0280822P.  
PR 02-APR-2001; 2001US-0280900P.  
PR 04-APR-2001; 2001US-0281194P.  
PR 13-APR-2001; 2001US-0283675P.  
PR 30-APR-2001; 2001US-0287424P.  
PR 02-MAY-2001; 2001US-0288066P.  
PR 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288528P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291199P.  
PR 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294889P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.  
PR 12-SEP-2001; 2001US-0318770P.  
PR 27-SEP-2001; 2001US-0325430P.  
PR 27-SEP-2001; 2001US-0325681P.  
PR 18-OCT-2001; 2001US-0330380P.

PR 31-OCT-2001; 2001US-0335301P.  
 PR 14-NOV-2001; 2001US-0332172P.  
 PR 14-NOV-2001; 2001US-0332271P.  
 PR 14-NOV-2001; 2001US-0332272P.  
 PR 14-NOV-2001; 2001US-0333184P.  
 PR 14-NOV-2001; 2001US-0333272P.  
 PR 21-NOV-2001; 2001US-0332094P.  
 PR 03-DEC-2001; 2001US-0337426P.  
 PR 03-DEC-2001; 2001US-0338032P.  
 PR 04-DEC-2001; 2001US-0337185P.  
 PR 03-JAN-2002; 2002US-0345705P.  
 PR 08-MAR-2002; 2002US-00093463.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ,  
 PI Boldog FI, Li L, Zetseren BD, Tchener VT, Gangolli EA, Vermet CAM;  
 PI Pena CEA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK,  
 PI Voss BZ, Malyankar UM, Anderson DW, Paturajan M, Miller CE;  
 PI Taulier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;  
 PI Zhong M;  
 XX  
 DR WPI; 2002-732824/79.  
 DR N-PSDB; ABV99422.  
 XX  
 PT New NOVX polypeptides and polynucleotides, useful for preventing,  
 PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,  
 PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic  
 PT disorders, and asthma.  
 XX  
 XX Claim 1; Page 264; 619pp; English.  
 XX  
 CC The present invention relates to new isolated proteins (NOVX) and their  
 CC coding sequences (ABV99327-ABV99595 and ABB70049-ABP70149), where X is  
 CC any number from 1 to 48. The NOVX proteins and coding sequences are  
 CC useful in the manufacture of a medicament for treating a syndrome  
 CC associated with a human disease, preferably a NOVX-associated disorder.  
 CC The NOVX coding sequences and proteins are useful for treating,  
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,  
 CC obesity, infectious disease, anorexia, cancer-associated cachexia,  
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's  
 CC disease, immune disorder, hematopoietic disorders, cardiovascular  
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic  
 CC disturbances associated with obesity, metabolic syndrome X or wasting  
 CC disorders associated with chronic diseases or various cancers. The NOVX  
 CC coding sequences and proteins may also be used as targets for the  
 CC identification of small molecules that modulate or inhibit e.g.  
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,  
 CC wound healing and angiogenesis, in gene therapy, in generation of  
 CC antibodies that bind immunospecifically to NOVX substances for use in  
 CC therapeutic or diagnostic methods  
 XX  
 SQ Sequence 778 AA;  
 Query Match 96.5%; Score 3219; DB 5; Length 778;  
 Best Local Similarity 96.1%; Pred. No. 1.4e-241;  
 Matches 617; Conservative 2; Mismatches 5; Indels 18; Gaps 1;  
 QY 1 MAPLLILLASGAACPVCVCONLSBSLSTCAHNGGLFVPPVNDRTVEYELADNFI 60  
 DB 1 MAPALLILLASGAACPVCVCONLSBSLSTCAHNGGLFVPPVNDRTVEYELADNFI 60  
 QY 61 QALGPPDFRNMGTGVDTLTSNATRTIGARFGLESRSRHLGNGRLVEIGTSLGSPV 120  
 DB 61 QALGPPDFRNMGTGVDTLTSNATRTIGARFGLESRSRHLGNGRLVEIGTSLGSPV 120  
 QY 121 NLGHLISGNQIGRIAGAPDFLESLEDLDSYNNLRQYVWAGIGAMPALHTLNDNL 180  
 DB 121 NLGHLISGNQIGRIAGAPDFLESLEDLDSYNNLRQYVWAGIGAMPALHTLNDNL 180  
 QY 181 IDALPAGAFQOLGSLRLDTLSNRLATLAPDPLFSRGRDAASAPVLSGNGPLHCNC 240  
 DB 181 IDALPAGAFQOLGSLRLDTLSNRLATLAPDPLFSRGRDAASAPVLSGNGPLHCNC 240

QY 241 ELLMLRLARPDDLETCASPPGAGRYFMAVPERGERSCEPPLIARHTQRLWLEGQATL 300  
 DB 241 ELLMLRLARPDDLETCASPPGAGRYFMAVPERGERSCEPPLIARHTQRLWLEGQATL 300  
 QY 301 RCRLGDPAPTMWVPPDRLVGNSSRARAFPNGTLEIGVTGAGDAGGYTCIATNPAGEA 360  
 DB 301 RCRLGDPAPTMWVPPDRLVGNSSRARAFPNGTLEIGVTGAGDAGGYTCIATNPAGEA 360  
 QY 361 TARVELRVIALPHGNSAEGGR-----PGPSDIASARTAAEGGT 402  
 DB 361 TARVELRVIALPHGNSAEGGRPCRTSAPMWEQCCRGCPSPSDIAASARTAAEGGT 420  
 QY 403 LSEEPVQVTEWATATGLVWGPGRPADPMMFOIYNSEDETLTYRIYPASSHFLTK 462  
 DB 421 LSEEPVQVTEWATATGLVWGPGRPADPMMFOIYNSEDETLTYRIYPASSHFLTK 480  
 QY 463 HLYPGADYDCLALSPAGPSDLTATRLGCAHFTLPAAPLCHALQAHVLTGLTVAV 522  
 DB 481 HLYPGADYDCLALSPAGPSDLTATRLGCAHFTLPAAPLCHALQAHVLTGLTVAV 540  
 QY 523 GGYLVNALLVFTYALLYRGRGAGNGRLPLKLSHVQSTNGGSPPTPKAHPRPSPRPOR 582  
 DB 541 GGYLVNALLVFTYALLYRGRGAGNGRLPLKLSHVQSTNGGSPPTPKAHPRPSPRPOR 600  
 QY 583 SCSLIDGADGCGYARPLGGAMARRSHSVHGILGACRGVG 624  
 DB 601 SCSLIDGADGCGYARPLGGAMARRSHSVHGILGACRGVG 642  
 RESULT 4  
 ADI21104  
 ID ADI21104 standard; protein; 618 AA.  
 AC ADI21104;  
 XX  
 DT 15-APR-2004 (first entry)  
 XX  
 DE Novel human protein #79.  
 XX  
 KM forensic; nutritional source; damaged tissue; diseased tissue;  
 KM myeloid cell disorder; lymphoid cell disorder;  
 KM bone cartilage tissue growth; tendon tissue growth;  
 KM ligament tissue growth; nerve tissue growth; regeneration; wound healing;  
 KM tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003025148-A2.  
 XX  
 PD 27-MAR-2003.  
 XX  
 PF 19-SEP-2002; 2002MO-US029964.  
 XX  
 PR 19-SEP-2001; 2001US-0323739P.  
 PR 13-SEP-2002; 2002US-00323739.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;  
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou F, Drmanac RT, Wang D;  
 PI Haley-Vicente D;  
 XX  
 DR WPI; 2003-354603/33.  
 DR N-PSDB; ADI21820.  
 XX  
 PT New polynucleotides and secreted proteins, useful for treating myeloid or  
 PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve  
 PT tissue growth or regeneration, in wound healing, and in tissue repair and  
 PT replacement.  
 XX  
 PS Claim 20; SEQ ID NO 355; 156pp; English.  
 XX

CC The invention relates to an isolated polynucleotide encoding a  
CC polypeptide with biological activity. The polynucleotides and  
CC polypeptides are useful in diagnostics, forensics, gene mapping,  
CC identification of mutations responsible for genetic disorders and other  
CC traits, to assess biodiversity, as nutritional sources or supplements.  
CC The polynucleotides may also be used as molecular weight markers,  
CC chromosome markers or map related gene positions, or as an antigen to  
CC raise anti-DNA antibodies or elicit immune response. The polypeptides are  
CC useful for raising antibodies, as markers for tissues in which the  
CC corresponding polypeptide is expressed, for re-engineering damaged or  
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in  
CC bone cartilage, tendon, ligament and/or nerve tissue growth or  
CC regeneration, in wound healing, in tissue repair and replacement, in  
CC healing of burns, incisions and ulcers, and in treating cancer. The  
CC present sequence represents the amino acid sequence of a novel human  
CC protein.  
CC  
XX  
XX

SQ Sequence 618 AA;

Query Match 96.5%; Score 3218.5; DB 7; Length 618;  
Best Local Similarity 97.0%; Pred. No. 1.1e-241;  
Matches 616; Conservative 0; Mismatches 2; Indels 17; Gaps 1;

QY 1 MAPPILLILLIAGAAACPLPCVCONTSESLSTCAHRLIFVPPNVDRTVEIRLADNFI 60  
DB 1 MAPPILLILLIAGAAACPLPCVCONTSESLSTCAHRLIFVPPNVDRTVEIRLADNFI 60  
QY 61 QALGPPDFRMTGLVLTLSRNAITRIGARFGDLSLSIHJDNRIYELGTSGIRGPV 120  
DB 61 QALGPPDFRMTGLVLTLSRNAITRIGARFGDLSLSIHJDNRIYELGTSGIRGPV 120  
QY 121 NIQHLLISGNQGRIPGAFDDFLESLEDLDSYNNLRQVPMAGIGAMPALHTLNDHNL 180  
DB 121 NIQHLLISGNQGRIPGAFDDFLESLEDLDSYNNLRQVPMAGIGAMPALHTLNDHNL 180  
QY 181 IDALPRGAPRQUGLSRLDITSNRLATLADPPIFSRGRDAEASPAVLVFSGNPLHCNC 240  
DB 181 IDALPRGAPRQUGLSRLDITSNRLATLADPPIFSRGRDAEASPAVLVFSGNPLHCNC 240  
QY 181 IDALPRGAPRQUGLSRLDITSNRLATLADPPIFSRGRDAEASPAVLVFSGNPLHCNC 240  
DB 181 IDALPRGAPRQUGLSRLDITSNRLATLADPPIFSRGRDAEASPAVLVFSGNPLHCNC 240  
QY 241 ELLMLRLARPDDLETCASPGLAGRYFMAVPEGSECEPPLIARHTORLMTVEGQRTL 300  
DB 241 ELLMLRLARPDDLETCASPGLAGRYFMAVPEGSECEPPLIARHTORLMTVEGQRTL 300  
QY 224 ELLMLRLARPDDLETCASPGLAGRYFMAVPEGSECEPPLIARHTORLMTVEGQRTL 283  
DB 224 ELLMLRLARPDDLETCASPGLAGRYFMAVPEGSECEPPLIARHTORLMTVEGQRTL 283  
QY 301 RCRLADPAFTMVMVGPDDRLVGNSSRABAFPMGTLEIGVTGADAGYTCIATNPAGBA 360  
DB 301 RCRLADPAFTMVMVGPDDRLVGNSSRABAFPMGTLEIGVTGADAGYTCIATNPAGBA 360  
QY 284 RCRLADPAFTMVMVGPDDRLVGNSSRABAFPMGTLEIGVTGADAGYTCIATNPAGBA 343  
DB 284 RCRLADPAFTMVMVGPDDRLVGNSSRABAFPMGTLEIGVTGADAGYTCIATNPAGBA 343  
QY 361 TARVELRVLALPRGNSABEGGPPSPDIAASARTAAEGGTLESEPAVQVTEVTATSGI 420  
DB 361 TARVELRVLALPRGNSABEGGPPSPDIAASARTAAEGGTLESEPAVQVTEVTATSGI 420  
QY 344 TARVELRVLALPRGNSABEGGPPSPDIAASARTAAEGGTLESEPAVQVTEVTATSGI 403  
DB 344 TARVELRVLALPRGNSABEGGPPSPDIAASARTAAEGGTLESEPAVQVTEVTATSGI 403  
QY 421 VSWGPRPADPVMVFOIYNSSEDTLIRIVPASHHPLKHLVPGADYDCLIALSPA 480  
DB 421 VSWGPRPADPVMVFOIYNSSEDTLIRIVPASHHPLKHLVPGADYDCLIALSPA 480  
QY 404 VSWGPRPADPVMVFOIYNSSEDTLIRIVPASHHPLKHLVPGADYDCLIALSPA 463  
DB 404 VSWGPRPADPVMVFOIYNSSEDTLIRIVPASHHPLKHLVPGADYDCLIALSPA 463  
QY 481 AGPSDLTATRLIGCAHFSTLPASPLCHALQAHVIGSTLIVAVGVVLAALLVFTVALVR 540  
DB 481 AGPSDLTATRLIGCAHFSTLPASPLCHALQAHVIGSTLIVAVGVVLAALLVFTVALVR 540  
QY 464 AGPSDLTATRLIGCAHFSTLPASPLCHALQAHVIGSTLIVAVGVVLAALLVFTVALVR 523  
DB 464 AGPSDLTATRLIGCAHFSTLPASPLCHALQAHVIGSTLIVAVGVVLAALLVFTVALVR 523  
QY 541 GRGAGGRPLPLKISHVQSTNGGPPPTPKAHPRRSPRRORSCTLDLADAGCYGARRI 600  
DB 541 GRGAGGRPLPLKISHVQSTNGGPPPTPKAHPRRSPRRORSCTLDLADAGCYGARRI 600  
QY 524 GRGAGGRPLPLKISHVQSTNGGPPPTPKAHPRRSPRRORSCTLDLADAGCYGARRI 583  
DB 524 GRGAGGRPLPLKISHVQSTNGGPPPTPKAHPRRSPRRORSCTLDLADAGCYGARRI 583  
QY 601 GGAMARRSHVHGGLIGAGRGVGSABRLSESVV 635  
DB 601 GGAMARRSHVHGGLIGAGRGVGSABRLSESVV 635  
QY 584 GGAMARRSHVHGGLIGAGRGVGSABRLSESVV 618  
DB 584 GGAMARRSHVHGGLIGAGRGVGSABRLSESVV 618

RESULT 5  
ID ABP70142  
XX ABP70142 standard; protein; 647 AA.  
AC ABP70142;

XX 27-JAN-2003 (first entry)  
XX Human NOV44a.  
XX Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;  
XX antineoplastic; cardiatic; haemostatic; neuroprotective; anorectic;  
XX neurotoxic; immunosuppressive; osteopathic; antiparkinsonian; cancer;  
XX antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;  
XX metabolic disorder; diabetes; obesity; infectious disease; anorexia;  
XX neurodegenerative disease; Alzheimer's disease; Parkinson's disease;  
XX immune disorder; haematopoietic disorder; cardiovascular disorder;  
XX bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;  
XX metabolic syndrome X; wasting disorder; cell differentiation;  
XX cell proliferation; haematopoiesis; wound healing; angiogenesis.  
XX  
OS Homo sapiens.  
XX  
XX WO200272771-A2.  
XX  
XX 19-SEP-2002.  
XX  
XX 08-MAR-2002; 2002MO-US007288.  
XX  
XX 08-MAR-2001; 2001US-0274101P.  
XX 08-MAR-2001; 2001US-0274194P.  
XX 08-MAR-2001; 2001US-0274281P.  
XX 08-MAR-2001; 2001US-0274322P.  
XX 09-MAR-2001; 2001US-0274849P.  
XX 12-MAR-2001; 2001US-0275355P.  
XX 13-MAR-2001; 2001US-0275578P.  
XX 13-MAR-2001; 2001US-0275579P.  
XX 13-MAR-2001; 2001US-0275601P.  
XX 14-MAR-2001; 2001US-0276000P.  
XX 16-MAR-2001; 2001US-027676P.  
XX 19-MAR-2001; 2001US-027694P.  
XX 20-MAR-2001; 2001US-0277239P.  
XX 20-MAR-2001; 2001US-0277321P.  
XX 20-MAR-2001; 2001US-0277327P.  
XX 20-MAR-2001; 2001US-0277327P.  
XX 21-MAR-2001; 2001US-0277791P.  
XX 22-MAR-2001; 2001US-0277833P.  
XX 23-MAR-2001; 2001US-0278152P.  
XX 26-MAR-2001; 2001US-0278894P.  
XX 27-MAR-2001; 2001US-0278999P.  
XX 27-MAR-2001; 2001US-0279036P.  
XX 28-MAR-2001; 2001US-0279344P.  
XX 30-MAR-2001; 2001US-0279959P.  
XX 30-MAR-2001; 2001US-0280233P.  
XX 02-APR-2001; 2001US-0280802P.  
XX 02-APR-2001; 2001US-0280802P.  
XX 02-APR-2001; 2001US-0280802P.  
XX 02-APR-2001; 2001US-0280802P.  
XX 04-APR-2001; 2001US-0281194P.  
XX 13-APR-2001; 2001US-0283675P.  
XX 30-APR-2001; 2001US-0287424P.  
XX 02-MAY-2001; 2001US-0288066P.  
XX 03-MAY-2001; 2001US-0288342P.  
XX 03-MAY-2001; 2001US-0288528P.  
XX 15-MAY-2001; 2001US-0291190P.  
XX 16-MAY-2001; 2001US-0291099P.  
XX 16-MAY-2001; 2001US-0291240P.  
XX 30-MAY-2001; 2001US-0294485P.  
XX 31-MAY-2001; 2001US-0294489P.  
XX 31-MAY-2001; 2001US-0294899P.  
XX 31-MAY-2001; 2001US-0294899P.  
XX 18-JUN-2001; 2001US-0299027P.  
XX 19-JUN-2001; 2001US-0299033P.  
XX 19-JUN-2001; 2001US-0299310P.  
XX 10-JUL-2001; 2001US-0304345P.  
XX 31-JUL-2001; 2001US-0309198P.  
XX 16-AUG-2001; 2001US-0313903P.  
XX 10-SEP-2001; 2001US-0318462P.  
XX 12-SEP-2001; 2001US-0318770P.  
XX 27-SEP-2001; 2001US-0325430P.



PR 27-SEP-2001; 2001US-0325681P.  
 PR 18-OCT-2001; 2001US-0330380P.  
 PR 31-OCT-2001; 2001US-0335301P.  
 PR 14-NOV-2001; 2001US-0332172P.  
 PR 14-NOV-2001; 2001US-0332271P.  
 PR 14-NOV-2001; 2001US-0332272P.  
 PR 14-NOV-2001; 2001US-0333184P.  
 PR 14-NOV-2001; 2001US-0333272P.  
 PR 21-NOV-2001; 2001US-0332094P.  
 PR 03-DEC-2001; 2001US-0337426P.  
 PR 03-DEC-2001; 2001US-0338092P.  
 PR 04-DEC-2001; 2001US-0337185P.  
 PR 03-JAN-2002; 2002US-0345705P.  
 PR 08-MAR-2002; 2002US-00093463.  
 PA (CURA-) CURAGEN CORP.  
 PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SC,  
 PI Boldog FI, Li L, Zernusen BD, Tcherven VT, Gangoli EA, Vernet CM,  
 PI Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK,  
 PI Voss EZ, Malyankar UM, Anderson DW, Paturajan M, Miller CE,  
 PI Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gueev VY, Pochart PF,  
 PI Zhong M,  
 DR WPI; 2002-732824/79.  
 DR N-PSDB; ABV99420.  
 XX  
 PT New NOVX polypeptides and polynucleotides, useful for preventing,  
 PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,  
 PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic  
 PT disorders, and asthma.  
 XX  
 PS Claim 1; Page 262; 619pp; English.  
 XX  
 CC The present invention relates to new isolated proteins (NOVX) and their  
 CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is  
 CC any number from 1 to 48. The NOVX proteins and coding sequences are  
 CC useful in the manufacture of a medicament for treating a syndrome  
 CC associated with a human disease, preferably a NOVX-associated disorder.  
 CC The NOVX coding sequences and proteins are useful for treating, diabetes,  
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,  
 CC obesity, infectious disease, anorexia, cancer-associated cachexia,  
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's  
 CC disease, immune disorders, hematopoietic disorders, cardiovascular  
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic  
 CC disturbances associated with obesity, metabolic syndrome X or wasting  
 CC disorders associated with chronic diseases or various cancers. The NOVX  
 CC coding sequences and proteins may also be used as targets for the  
 CC identification of small molecules that modulate or inhibit e.g.  
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,  
 CC wound healing and angiogenesis, in gene therapy, in generation of  
 CC antibodies that bind immunospecifically to NOVX substances for use in  
 CC therapeutic or diagnostic methods  
 XX  
 SQ Sequence 647 AA;  
 Query March 90.7%; Score 3026.5; DB 5; Length 647;  
 Best Local Similarity 93.9%; Pred. No. 1.1e-226;  
 Matches 589; Conservative 0; Mismatches 21; Indels 17; Gaps 3;  
 QY 1 MAPRLILLILASGAACPLPCVCONLSLSLTLCAHRLGLFVPPNVDRRTVELRLADNFI 60  
 DB 1 MAPRLILLILASGAACPLPCVCONLSLSLTLCAHRLGLFVPPNVDRRTVELRLADNFI 60  
 QY 61 QALGPPFRMTGIVDTLTSRNATRTTGARAFGLSELSRLHLDGNLVEIGTSLGFPV 120  
 DB 61 QALGPPFRMTGIVDTLTSRNATRTTGARAFGLSELSRLHLDGNLVEIGTSLGFPV 120  
 QY 121 NLGHLISGQOLGRIAGAFDDFLESLEDDLSTYNNROYVMAGIGAMPALHTLNLHNL 180  
 DB 121 NLGHLISGQOLGRIAGAFDDFLESLEDDLSTYNNROYVMAGIGAMPALHTLNLHNL 180  
 QY 181 IDALPPGAFAQLGQLSRDLTNSRLATLAADPLFRSGRDAEAPAPLVLSFGNPLHCNC 240

DB 181 IDALPPGAFAQLGQLSRDLTNSRLATLAADPLFRSGRDAEAPAPLVLSFGNPLHCNC 240  
 QY 241 ELLMLRLARPDDLETCASPGLAGRYFWAVPGEFSCEPPLTARHTORLVLBEGORATL 300  
 DB 241 ELLMLRLARPDDLETCASPGLAGRYFWAVPGEFSCEPPLTARHTORLVLBEGORATL 300  
 QY 301 RCRAIGDPAFTMMWVGDDDLVGNSSRARAFEPGTLTIGVTAGDAGGYTCIATNPAGEA 360  
 DB 301 RCRAIGDPAFTMMWVGDDDLVGNSSRARAFEPGTLTIGVTAGDAGGYTCIATNPAGEA 360  
 QY 361 TARVELRVLALPHGNGSSAAGRPGPSDIAASARTAAEGGTLSSEBAVQTEVTATSGL 420  
 DB 361 TARVELRVLALPHGNGSSAAGRPGPSDIAASARTAAEGGTLSSEBAVQTEVTATSGL 420  
 QY 421 VSWGPRPADPVMFQIQVNSSEDETLIRIVASSHHFLKXVLVQADYDLCLIALSPA 480  
 DB 421 VSWGPRPADPVMFQIQVNSSEDETLIRIVASSHHFLKXVLVQADYDLCLIALSPA 480  
 QY 481 AGPSDLTATRLGCAHSTLPASPLCHALQAHVGLGTLTAVAGGYVAALLVFTVALLVR 540  
 DB 481 AGPSDLTATRLGCAHSTLPASPLCHALQAHVGLGTLTAVAGGYVAALLVFTVALLVR 540  
 QY 541 GRGAGNGRLPLKLSHVOSQTNQSPSPTPKXAPPRSP-----PRPORSCTLDLDAG 592  
 DB 541 GRGAGNGRLPLKLSHVOSQTNQSPSPTPKXAPPRSP-----PRPORSCTLDLDAG 592  
 QY 593 CYGYARRLGG-----AMARSHVHGG 614  
 DB 593 CYGYARRLGG-----AMARSHVHGG 614  
 QY 597 CQAPGRSLGPTBPLCAMGAARGVPGG 623  
 DB 597 CQAPGRSLGPTBPLCAMGAARGVPGG 623  
 RESULT 6  
 ADE08361  
 ID ADE08361 standard; protein; 565 AA.  
 AC ADE08361;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 XX Novel protein (useful for identifying genetic disorders) #516.  
 XX  
 XX novel gene; novel protein; tissue marker; molecular weight marker;  
 XX chromosome marker; genetic disorder.  
 OS Unidentified.  
 XX  
 XX WC02003054152-A2.  
 PN  
 PD 03-JUL-2003.  
 XX  
 PF 10-DEC-2002; 2002WC-US039555.  
 XX  
 PR 10-DEC-2001; 2001US-0339739P.  
 PR 11-DEC-2001; 2001US-0339453P.  
 PR 14-MAR-2002; 2002US-0365091P.  
 PR 14-MAR-2002; 2002US-0365384P.  
 PR 12-APR-2002; 2002US-0372381P.  
 PR 12-APR-2002; 2002US-0372615P.  
 PR 22-APR-2002; 2002US-0012855P.  
 PR 24-APR-2002; 2002US-0376045P.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Auanti V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J,  
 PI Ghosh M, Xue AD, Wehman T, Weng G, Zhou F, Drmanac RT, Wang Z;  
 PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;  
 XX  
 DR WPI; 2003-569235/53.  
 DR N-PSDB; ADE07450.  
 XX  
 PT New polynucleotides, useful for expressing recombinant proteins for  
 PT analysis, characterization or therapeutic use, or as markers for tissues

PT in which the corresponding protein is preferentially expressed.  
XX Claim 20; SEQ ID NO 1427; 1177pp; English.  
XX  
CC The invention comprises the amino acid and coding sequences of novel  
CC proteins. The DNA and protein sequences of the invention are useful as:  
CC markers for tissues in which the corresponding protein is preferentially  
CC expressed; as molecular weight markers on gels; as chromosome markers or  
CC tags; to identify chromosomes or to map related gene positions; and to  
CC compare with endogenous DNA sequences in patients to identify potential  
CC genetic disorders. The present amino acid sequence represents a protein  
CC of the invention.  
XX  
SQ Sequence 565 AA:  
  
Query Match 89.0%; Score 2970; DB 7; Length 565;  
Best Local Similarity 100.0%; Pred. No. 2, 2e-222;  
Matches 565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 71 MTGLVLTLSRNAITRIGARAFGDLSELSRLHLDGNRLVELTGSIRGPVNLQHLILSGN 130  
Db 1 MTGLVLTLSRNAITRIGARAFGDLSELSRLHLDGNRLVELTGSIRGPVNLQHLILSGN 60  
QY 131 QUGRIAPGAFDDLELELDLSSYNMLROYPMAGIGAMPALHTLNDHNLIDALPGARA 190  
Db 61 QUGRIAPGAFDDLELELDLSSYNMLROYPMAGIGAMPALHTLNDHNLIDALPGARA 120  
QY 191 QUGLSRLDLTSNRALTAPDPLFSRGRDAEAPALVLSFGSNPLHNCCELLMLRLAR 250  
Db 121 QUGLSRLDLTSNRALTAPDPLFSRGRDAEAPALVLSFGSNPLHNCCELLMLRLAR 180  
QY 251 PDDLETCASPPGIAGIYFMAVPEGEFSCPEPLIARTTORLWLEGGORATLRCALGDPAP 310  
Db 181 PDDLETCASPPGIAGIYFMAVPEGEFSCPEPLIARTTORLWLEGGORATLRCALGDPAP 240  
QY 311 TMMWGPDDRLVNSSRRARAFPGTLEIGYTGAGDAGCTTCTATNPAGATARVELRVIA 370  
Db 241 TMMWGPDDRLVNSSRRARAFPGTLEIGYTGAGDAGCTTCTATNPAGATARVELRVIA 300  
QY 371 LPHGNSASAGRPGRPSDIAASARTAAEGEGLSESPAVQVTEVATSGLVSGPGRPAD 430  
Db 301 LPHGNSASAGRPGRPSDIAASARTAAEGEGLSESPAVQVTEVATSGLVSGPGRPAD 360  
QY 431 PVMWFOIQYNSSEDETLIRIVPASHHFLKHLVPGADYDCLIALSPAPSPDLTATR 490  
Db 361 PVMWFOIQYNSSEDETLIRIVPASHHFLKHLVPGADYDCLIALSPAPSPDLTATR 420  
QY 491 ILGCAHFTLPLASPLCHALQAHVIGSTLVAVGVVALLVFTVALIVRGAGNGRLP 550  
Db 421 ILGCAHFTLPLASPLCHALQAHVIGSTLVAVGVVALLVFTVALIVRGAGNGRLP 480  
QY 551 IMLSHVQQTNGSPPTKAPHPRRSPRRSCSIDLDGACGYARLGGAMARRSHS 610  
Db 481 IMLSHVQQTNGSPPTKAPHPRRSPRRSCSIDLDGACGYARLGGAMARRSHS 540  
QY 611 VHGGILGAGCGVGGSAERLESVV 635  
Db 541 VHGGILGAGCGVGGSAERLESVV 565  
  
RESULT 7  
ABP70143  
ID ABP70143 standard; protein; 566 AA.  
XX  
XX ABP70143;  
AC  
XX  
XX 27-JAN-2003 (first entry)  
DT  
XX  
XX Human NOV44b.  
DE  
XX  
XX Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;  
KW antiinflammatory; cardiostatic; haemostatic; neuroprotective; anorectic;  
KW neurotropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;

KW antiinfertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;  
KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;  
KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; hematopoietic disorder; cardiovascular disorder;  
KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;  
KW bronchial syndrome X; wasting disorder; cell differentiation;  
KW cell proliferation; hematopoiesis; wound healing; angiogenesis.  
OS  
XX Homo sapiens.  
PN WO200272771-A2.  
PD  
XX  
PD 19-SEP-2002.  
XX  
PF 08-MAR-2002; 2002WO-US007288.  
XX  
PR 08-MAR-2001; 2001US-0274101P.  
XX  
PR 08-MAR-2001; 2001US-0274194P.  
PR 08-MAR-2001; 2001US-0274281P.  
PR 08-MAR-2001; 2001US-0274322P.  
PR 09-MAR-2001; 2001US-0274849P.  
PR 12-MAR-2001; 2001US-0275235P.  
PR 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
PR 13-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
PR 16-MAR-2001; 2001US-0276776P.  
PR 19-MAR-2001; 2001US-0276994P.  
PR 20-MAR-2001; 2001US-0277239P.  
PR 20-MAR-2001; 2001US-0277321P.  
PR 20-MAR-2001; 2001US-0277327P.  
PR 20-MAR-2001; 2001US-0277381P.  
PR 21-MAR-2001; 2001US-0277791P.  
PR 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
PR 26-MAR-2001; 2001US-0278894P.  
PR 27-MAR-2001; 2001US-0278999P.  
PR 27-MAR-2001; 2001US-0279036P.  
PR 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0279995P.  
PR 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280802P.  
PR 02-APR-2001; 2001US-0280822P.  
PR 02-APR-2001; 2001US-0280900P.  
PR 04-APR-2001; 2001US-0281194P.  
PR 13-APR-2001; 2001US-0283675P.  
PR 30-APR-2001; 2001US-0287424P.  
PR 02-MAY-2001; 2001US-0288066P.  
PR 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288528P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
PR 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.  
PR 12-SEP-2001; 2001US-0318770P.  
PR 27-SEP-2001; 2001US-0325430P.  
PR 27-SEP-2001; 2001US-0325681P.  
PR 18-OCT-2001; 2001US-0330380P.  
PR 31-OCT-2001; 2001US-0335301P.  
PR 14-NOV-2001; 2001US-0332172P.  
PR 14-NOV-2001; 2001US-0332271P.  
PR 14-NOV-2001; 2001US-0332272P.  
PR 14-NOV-2001; 2001US-0333184P.  
PR 14-NOV-2001; 2001US-0333272P.

PR	21-NOV-2001;	2001US-0332094P.
PR	03-DEC-2001;	2001US-0337426P.
PR	03-DEC-2001;	2001US-0338092P.
PR	04-DEC-2001;	2001US-0337185P.
PR	03-JAN-2002;	2002US-0345705P.
PR	08-MAR-2002;	2002US-00093463.
XX		
XX		(CURA-) CURAGEN CORP.
PA		
PI	Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;	
PI	Boldo FL, Li L, Zettnusen BD, Tcherny VT, Gangoli RA, Vernet CM;	
PI	Pena CEA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;	
PI	Voss EZ, Malyarkar UM, Anderson DW, Patluraan M, Miller CE;	
PI	Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;	
PI	Zhong M;	
XX		
XX		WPI: 2002-773824/79.
DR		N-PSDB; ABV99421.
XX		
PT	New NOXV polypeptides and polynucleotides, useful for preventing,	
PT	diagnosing or treating NOXV-associated disorders e.g. diabetes, cancer,	
PT	Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic	
PT	disorders, and asthma.	
XX		
XX		Claim 1; Page 263; 619pp; English.
XX		
CC	The present invention relates to new isolated proteins (NOXV) and their	
CC	coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is	
CC	any number from 1 to 48. The NOXV proteins and coding sequences are	
CC	useful in the manufacture of a medicament for treating a syndrome	
CC	associated with a human disease, preferably a NOXV-associated disorder.	
CC	The NOXV coding sequences and proteins are useful for treating,	
CC	preventing or diagnosing diseases such as metabolic disorders, diabetes,	
CC	obesity, infectious disease, anorexia, cancer-associated cachexia,	
CC	cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's	
CC	disease, immune disorders, haematopoietic disorders, cardiovascular	
CC	disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic	
CC	disturbances associated with obesity, metabolic syndrome X or wasting	
CC	disorders associated with chronic diseases or various cancers. The NOXV	
CC	coding sequences and proteins may also be used as targets for the	
CC	identification of small molecules that modulate or inhibit e.g.	
CC	neurogenesis, cell differentiation, cell proliferation, haematopoiesis,	
CC	wound healing and angiogenesis, in gene therapy, in generation of	
CC	antibodies that bind immunospecifically to NOXV substances for use in	
CC	therapeutic or diagnostic methods	
XX		
SQ	Sequence 566 AA;	
	Query Match 83.1%; Score 2773.5; DB 5; Length 566;	
	Best Local Similarity 95.4%; Pred. No. 4, 5e-207;	
	Matches 537; Conservative 4; Mismatches 11; Indels 11; Gaps 2	
QY	1 MAPPLILLILLASGAACPLPCYCCNLSSSLTCAHRLGLFPVPVDRRTVRLADNPI 60	
Db	1 MAPPLILLILLASGAACPLPCVCCNLSSSLTCAHRLGLFPVPVDRRTVRLADNPI 60	
QY	61 QALGPDPERNMTGLVDLTLSRNAIRIRIGARAFGDELSRLSLDGNRLVELTGSIRGEV 120	
Db	61 QALGPDPERNMTGLVDLTLSRNAIRIRIGARAFGDELSRLSLDGNRLVELTGSIRGEV 120	
QY	121 NLQHLILSGNOLGRIAPGAFDDFLSLLEDLDLSYNNLROVPVPAAGIGAMPALTLINDHL 180	
Db	121 NLQHLILSGNOLGRIAPGAFDDFLSLLEDLDLSYNNLROVPVPAAGIGAMPALTLINDHL 180	
QY	181 IDALPFGFAQOLGQSLRLDLSNRRLATLAPDLFSGRPAEASPAFLVLSFGSNPLHNC 240	
Db	181 IDALPFGFAQOLGQSLRLDLSNRRLATLAPDLFSGRPAEASPAFLVLSFGSNPLHNC 240	
QY	241 ELLWRLRLARPDDLTCASPGLAGRYFWAVGEFGSCPPIIARHTQRLWLVEGORATL 300	
Db	241 ELLWRLRLARPDDLTCASPGLAGRYFWAVGEFGSCPPIIARHTQRLWLVEGORATL 300	
QY	301 RCRALGDAFPTMHWGVPDDRLVGNSSRAAFVNGTLEIVTGAGDAGGTYCTIATNPGA 360	

Db 301 RCALADDPAPTHWVPPDDRLVWNSRRAPENGLTEGVIGADAGGYTCIATNPGEA 360  
 Qy 361 TARVELCRVALPHGGNSSAEGGRPPSPDIASARTAAEGEETLESEPAVQYTEVATISGL 420  
 Db 361 TARVELRVIALPHGGNSSAEGGRPPSPDIASARTAAEGEETLESEPAVQYTEVATISGL 420  
 Qy 421 VSMGPRRPAPDPMMWQIQYNSSSEDTLYRIVPASSHHFLKHHVPGADYDCLIALSPA 480  
 Db 421 VSMGPRRPAPDPMMWQIQYNSSSEDTLYRIVPASSHHFLKHHVPGADYDCLIALSPA 480  
 Qy 481 AGRSDDTATRLIGCAHFSITLPASPLCHALQAHVLGGITVTVAAGVILVAALLVFTVALVR 540  
 Db 481 AGRSDDTATRLIGCAHFSITLPASPLCHALQAHVLGGITVTVAAGVILVAALPA-SVAVSTC 539  
 Qy 541 GRG-----AGNGRLPLKL 553  
 Db 540 WRGLPLPMGGQNRKGSLPLQV 562

CC	growth, including retinoblastoma, renal cell adenocarcinoma, endometrial
CC	useful in diagnosing and treating disorders related to abnormal cell
CC	proteins in general from cells and tissues. The Zlrr DNA and proteins are
CC	proteins of interest from a host cell and to monitor the secretion of
CC	Zlrr7, Zlrr8, and Zlrr9 proteins are useful in directing the secretion of
CC	repeat proteins Zlrr7, Zlrr8, Zlrr9 and spliced variants of Zlrr8, Zlrr9
CC	The invention relates to an isolated polypeptide comprising leucine-rich
XX	Claim 10; Page 67-69; 82pp; English.
XX	
PT	New leucine-rich repeat proteins and polymucleotides, useful for
PT	diagnosing and treating disorders related to abnormal cell growth e.g.
PT	retinoblastoma, renal cell, endometrial adenocarcinoma, leukemia, kidney
PT	and lung tumors.
XX	
XX	
XX	
XX	Thayer EC, Sheppard PO, Presnell SR;
PI	
XX	
XX	WPI: 2002-154725/20.
DR	N-PSDB; AAD28124.
DR	
XX	
PA	(ZYMO ) ZYMOGENETICS INC.
XX	
XX	30-JUN-2000; 2000US-0215446P.
PR	
XX	
XX	02-JUL-2001; 2001WO-US020999.
PF	
XX	
PD	10-JAN-2002.
XX	
PN	WO200202604-A2.
XX	
FT	Protein /label= Mature_Zlrr8_protein
FT	19.551
FT	/label= Signal_peptide
FT	1.18
FH	Key Location/Qualifiers
XX	
OS	Homo sapiens.
XX	
KW	rhabdomyosarcoma; genitourinary tract; chromosome 11q13.
KW	mammary gland; germ cell; bladder; oesophagus; pancreas; animal feed;
KW	cell growth disorder; glioblastoma; neuroblastoma; kidney tumour; lung;
KM	endometrial adenocarcinoma; renal cell; colon; prostate; retinoblastoma;
KM	Human; leucine-rich repeat-8; ZLRR8; cytostatic; gene therapy; leukaemia;
XX	
DE	Human leucine-rich repeat-8 (ZLRR8) protein #2.
XX	
DJ	22-APR-2002 (first entry)
XX	
AC	AAE17484;
XX	
ID	AAE17484 standard; protein; 551 AA.
RESULT 8	
AAE17484	

CC adenocarcinoma, glioblastoma, neuroblastoma, B-cell lymphocytic leukaemia,  
 CC kidney tumours, germ cell tumours, lung large cell carcinoma, mammary,  
 CC colon adenocarcinoma, genitourinary tract transitional cell tumours,  
 CC rhabdomyosarcoma, lung tumour, bladder tumour, oesophagus, pancreas and  
 CC and prostate adenocarcinoma. Zlrr protein is useful for identifying agonists  
 CC and antagonists of the polypeptide, for drug design, to screen for cell  
 CC metabolism affecting receptors, for analysis of cell phenotype, and as  
 CC animal feed supplement and cell culture components. Zlrr DNA is also  
 CC useful in gene therapy. The present sequence is human Zlrr protein.  
 CC Zlrr gene is located on chromosome 11q13  
 CC  
 XX  
 SO Sequence 551 AA;

Query Match 83.1%; Score 2773; DB 5; Length 551;  
 Best Local Similarity 100.0%; Pred. No. 4,8e-207;  
 Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPPILLILLASGAAACPLPCVCONLSSESLSTLCAHRLGLFVPPNVDRTVELRLADNFI 60  
 DB 1 MAPPILLILLASGAAACPLPCVCONLSSESLSTLCAHRLGLFVPPNVDRTVELRLADNFI 60  
 QY QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHDGNRLVGLGSGRGPV 120  
 DB QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHDGNRLVGLGSGRGPV 120  
 QY 121 NIQHLILSGNQLGRIARAGAFDFLESLEDLDSYNNLRQVPMAGIGAMPALHTLNDHN 180  
 DB 121 NIQHLILSGNQLGRIARAGAFDFLESLEDLDSYNNLRQVPMAGIGAMPALHTLNDHN 180  
 QY 181 IDLPPGAFQOLGSLRDLTNSRLATLADPLFSRGRDAEAPVLVSFGNPLHCNC 240  
 DB 181 IDLPPGAFQOLGSLRDLTNSRLATLADPLFSRGRDAEAPVLVSFGNPLHCNC 240  
 QY 241 ELIMRLRLAPDDLETCAAPGLAGRYFMAVPEGEFSCPEPLIARHQRVMEVSGRATL 300  
 DB 241 ELIMRLRLAPDDLETCAAPGLAGRYFMAVPEGEFSCPEPLIARHQRVMEVSGRATL 300  
 QY 301 ELMRLRLAPDDLETCAAPGLAGRYFMAVPEGEFSCPEPLIARHQRVMEVSGRATL 360  
 DB 301 ELMRLRLAPDDLETCAAPGLAGRYFMAVPEGEFSCPEPLIARHQRVMEVSGRATL 360  
 QY 361 RCALGDPAPFTMHWGPDRLVGNSSRRARFPNGTLEIGVTGAGDAGYTCIATNPAGEA 420  
 DB 361 RCALGDPAPFTMHWGPDRLVGNSSRRARFPNGTLEIGVTGAGDAGYTCIATNPAGEA 420  
 QY 421 VSWGPPRPADPVMWFOIQYNSSEDETLIYRIVPASSHHFLKHLVPGADVDLCLALSPA 480  
 DB 421 VSWGPPRPADPVMWFOIQYNSSEDETLIYRIVPASSHHFLKHLVPGADVDLCLALSPA 480  
 QY 481 AGPSDLTARRLTGCAHFSTLPASPLCHALQAHVLGGTLTVAVGVVAAL 530  
 DB 481 AGPSDLTARRLTGCAHFSTLPASPLCHALQAHVLGGTLTVAVGVVAAL 530

## RESULT 9

ABG04827  
 ID ABG04827 standard; protein; 526 AA.

AC ABG04827;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #4818.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 XX 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 PI Dermanac RT, Liu C, Tang YT;  
 XX  
 XX WPI, 2001-639362/73.  
 DR N-PSDB; AAS69014.  
 XX  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 XX Claim 20; SEQ ID NO 35186; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 XX

Sequence 526 AA;

Query Match 77.0%; Score 2568.5; DB 4; Length 526;  
 Best Local Similarity 89.3%; Pred. No. 3,8e-191;  
 Matches 502; Conservative 4; Mismatches 15; Indels 41; Gaps 2;

QY 1 MAPPILLILLASGAAACPLPCVCONLSSESLSTLCAHRLGLFVPPNVDRTVELRLADNFI 60  
 DB 1 MAPPILLILLASGAAACPLPCVCONLSSESLSTLCAHRLGLFVPPNVDRTVELRLADNFI 60  
 QY QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHDGNRLVGLGSGRGPV 119  
 DB QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHDGNRLVGLGSGRGPV 119  
 QY 61 QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHDGNRLVGLGSGRGPV 120  
 DB 61 QALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLSELSLHDGNRLVGLGSGRGPV 120  
 QY 120 VNIQHLILSGNQLGRIARAGAFDFLESLEDLDSYNNLRQVPMAGIGAMPALHTLNDHN 179  
 DB 120 VNIQHLILSGNQLGRIARAGAFDFLESLEDLDSYNNLRQVPMAGIGAMPALHTLNDHN 180  
 QY 181 IDLPPGAFQOLGSLRDLTNSRLATLADPLFSRGRDAEAPVLVSFGNPLHCNC 240  
 DB 181 IDLPPGAFQOLGSLRDLTNSRLATLADPLFSRGRDAEAPVLVSFGNPLHCNC 240  
 QY 240 CELLMRLRLAPDDLETCAAPGLAGRYFMAVPEGEFSCPEPLIARHQRVMEVSGRAT 299  
 DB 240 CELLMRLRLAPDDLETCAAPGLAGRYFMAVPEGEFSCPEPLIARHQRVMEVSGRAT 300  
 QY 300 LRCRALGDPAPFTMHWGPDRLVGNSSRRARFPNGTLEIGVTGAGDAGYTCIATNPAGE 359  
 DB 300 LRCRALGDPAPFTMHWGPDRLVGNSSRRARFPNGTLEIGVTGAGDAGYTCIATNPAGE 360  
 QY 360 ATARVRLRLALPHGNSSAEGGRPPSDIAASARTAEGBGLESEPAVQVTEVATISG 419  
 DB 360 ATARVRLRLALPHGNSSAEGGRPPSDIAASARTAEGBGLESEPAVQVTEVATISG 419

Db 361 ATARVELRVLALPHGNSAEGRRP----- 387

QY 420 LVSNGPRPADPVMFOIQNSSEDETLIYRIVPASSHHFLKHLVPGADYDCLIALSP 479

Db 388 -----RTSPPLALPRVSSSEDETLIYRIVPASSHHFLKHLVPGADYDCLIALSP 440

QY 480 AAGPSDITARILGCAHSTLPASPLCHALQAHVIGTLTVAVGCVVAALVFTVALV 539

Db 441 AAGSDITARILGCAHSTLPASPLCHALQAHVIGTLTVAVGCVVAALVFTVALV 500

QY 540 RGRGAGNGRLPLKLSHVOSQTN 561

Db 501 RGRGAGNGRLPLKLSHVOSQTN 522

RESULT 10

ID AAB70072 standard; protein; 468 AA.

XX AAB70072;

AC AAB70072;

XX 14-MAY-2001 (first entry)

XX Human secreted protein #11.

DE Human secreted protein #11.

XX Human; secreted protein; immunomodulatory; antisclerotic; dermatological;

KM immunosuppressive; antineoplastic; anti-HIV; immunostimulant;

KM cytostatic; cardiac; vascular; anti-angiogenic; ophthalmological;

KM neuroprotectant; anticonvulsant; vaccine; antialzheimer's;

KM antiparkinsonian; antimicrobial; vulnery; gene therapy;

KM immune disorder; hyperproliferative; cardiovascular; angiogenic;

KM neurological; infection.

XX

OS Homo sapiens.

XX WO200112776-A2.

PN 22-FEB-2001.

PD 15-AUG-2000; 2000WO-US022350.

PF 16-AUG-1999; 99US-0148759P.

XX

XX (HUMA-) HUMAN GENOME SCI INC.

PA Shu Y, Young PE, Ebner R, Soppet DR, Ruben SM;

PI WPI; 2001-244245/25.

DR N-PSDB; AAF76853.

XX

PT Nucleic acids encoding 18 human secreted polypeptides, useful for

PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease

XX and diabetic retinopathy.

XX Claim 11; Page 367-368; 380bp; English.

XX

CC The present sequence is one of 18 novel human secreted proteins. The

CC nucleic acids encoding the proteins and the proteins themselves may be

CC used in the prevention, diagnosis and treatment of diseases including

CC immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus

CC and human immunodeficiency virus (HIV) infections), hyperproliferative

CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases

CC (e.g. Schmitz syndrome, Chaga's cardiomyopathy and coronary

CC arteriosclerosis), angiogenic disorders (e.g. corneal graft

CC neovascularization and diabetic retinopathy), neurological disorders

CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),

CC infectious diseases and/or for promoting wound healing, regeneration

CC and/or chemotaxis. The nucleic acid molecules may also be used as DNA

CC probes in diagnostic assays to detect and quantitate the presence of

CC similar nucleic acid sequences in samples. The polypeptides may also be

CC used as antigens in the production of antibodies and in assays to

CC identify modulators of protein expression and activity

XX

SEQ Sequence 468 AA:

Query Match 74.2%; Score 2475; DB 4; Length 468;

Best Local Similarity 99.8%; Pred. No. 6.3e-184;

Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 168 MPALHTLNDPHNIDLPFGAFQOLGSLRDLTSNRILATLADPPFSRGRDSEASAPL 227

Db 1 MPALHTLNDPHNIDLPFGAFQOLGSLRDLTSNRILATLADPPFSRGRDSEASAPL 60

QY 228 VLSFSGNPLHNCCELLMLRLARPDDLETCASPPGLAGRYFMAVPEGEFSCPEPLIARHT 287

Db 61 VLSFSGNPLHNCCELLMLRLARPDDLETCASPPGLAGRYFMAVPEGEFSCPEPLIARHT 120

QY 288 QRLWVEGGRATRCALGDPATMHWGDDRLVGNSSARAPNGLTLEIGATGAGDAG 347

Db 121 QRLWVEGGRATRCALGDPATMHWGDDRLVGNSSARAPNGLTLEIGATGAGDAG 180

QY 348 GYTICATNPAGEATARVELRVLALPHGNSAEGRRPSPDIAASARTAAEGGTLESEP 407

Db 181 GYTICATNPAGEATARVELRVLALPHGNSAEGRRPSPDIAASARTAAEGGTLESEP 240

QY 408 AVGVTEVTATSGLVSWGPRPADPVMFOIQNSSEDETLIYRIVPASSHHFLKHLVPG 467

Db 241 AVGVTEVTATSGLVSWGPRPADPVMFOIQNSSEDETLIYRIVPASSHHFLKHLVPG 300

QY 468 ADVDLCLIALSPAGPSDLTATRLGCAHSTLPASPLCHALQAHVIGTLTVAVGCVLV 527

Db 301 ADVDLCLIALSPAGPSDLTATRLGCAHSTLPASPLCHALQAHVIGTLTVAVGCVLV 360

QY 528 AALLVFTVALLVGRGANGRLPLKLSHVOSQTNCGSPPTPKAHPSPPPRQSCSLD 587

Db 361 AALLVFTVALLVGRGANGRLPLKLSHVOSQTNCGSPPTPKAHPSPPPRQSCSLD 420

QY 588 LGDAGCYGARRLGAMARRSHVHGILLGACRGVGSABRLBESEV 635

Db 421 LGDAGCYGARRLGAMARRSHVHGILLGACRGVGSABRLBESEV 468

RESULT 11

ID AAG65511 standard; protein; 468 AA.

XX AAG65511;

AC AAG65511;

XX 27-AUG-2002 (first entry)

XX

XX Human albumin fusion protein #2186.

DE

XX

XX Albumin fusion protein; therapeutic protein X; human albumin; HA;

KM human serum albumin; HSA; cancer; reproductive disorder;

KM digestive disorder; immune disorder; endocrine disorder;

KM haematopoietic disorder; neural disorder; connective disorder;

KM cyostatic; antifertility; antineoplastic; antitumor;

KM immunomodulatory; anti-HIV; antidiabetic; haemostatic; noctropic;

KM neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;

KM osteopathic; antiarthritic.

XX

OS Homo sapiens.

XX Synthetic.

OS WO200177137-A1.

PN 18-OCT-2001.

XX

PD 12-APR-2001; 2001WO-US011988.

XX

XX 12-APR-2000; 2000US-0229358P.

XX 25-APR-2000; 2000US-0199384P.

PR 21-DEC-2000; 2000US-0256931P.

XX

XX (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Haseltine WA;  
 XX WPI; 2002-010886/01.  
 XX New fusion protein for treating disease e.g. diabetes comprises an  
 PT albumin fused to a therapeutic protein.  
 PS Claim 1; Page 2080-2081; 2102pp; English.  
 CC The present invention relates to albumin fusion proteins comprising a  
 CC therapeutic protein X and human albumin (HA, also known as human serum  
 CC albumin, HSA). The proteins are useful for treating a disease or disorder  
 CC that may be modulated by therapeutic protein X. The albumin extends the  
 CC shelf-life of protein X, and may increase its biological in vitro/in vivo  
 CC activity. The protein is useful for treating and diagnosing disorders  
 CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's  
 CC disease, ulcerative colitis), immune disorders (e.g. acquired  
 CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),  
 CC hematopoietic disorders, neural disorders (e.g. Alzheimer's,  
 CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,  
 CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).  
 CC ABG63326-ABG65518 represent albumin fusion proteins of the invention  
 XX  
 SQ Sequence 468 AA;

Query Match 74.2%; Score 2475; DB 5; Length 468;  
 Best Local Similarity 99.8%; Pred. No. 6.3e-184;  
 Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 168 MPALHTLNDHNLIDALPPGAFQGLGSLRLDLSNRRLTLADPLFSGRDAEAPAPL 227  
 DB 1 MPALHTLNDHNLIDALPPGAFQGLGSLRLDLSNRRLTLADPLFSGRDAEAPAPL 60  
 QY 228 VLSFGSNPLHNCCELLMLRLARPDDLTCASPPGAGRYFMAVPEGESCEPPLFARHT 287  
 DB 61 VLSFGSNPLHNCCELLMLRLARPDDLTCASPPGAGRYFMAVPEGESCEPPLFARHT 120  
 QY 288 QRLWVEGQRATLRCAAGDPAPYTMWVGDDRLVGNSSRAAPFPNGTLEIGVAGADAG 347  
 DB 121 QRLWVEGQRATLRCAAGDPAPYTMWVGDDRLVGNSSRAAPFPNGTLEIGVAGADAG 180  
 QY 348 GTCTCIANTPAGETAVEIRVIALPPGNSAEGGRPGSDTAAKRTAAEGEGTLESBP 407  
 DB 181 GTCTCIANTPAGETAVEIRVIALPPGNSAEGGRPGSDTAAKRTAAEGEGTLESBP 240  
 QY 408 AVQVTEVTATSGVSWGPRPADPYWMFOIQYNSSEDETLIRIYVAPASHHFLKHLVVG 467  
 DB 241 AVQVTEVTATSGVSWGPRPADPYWMFOIQYNSSEDETLIRIYVAPASHHFLKHLVVG 300  
 QY 468 AVYDCLALSPAGPSDLTATRLGCAHFSTLPASPLCHALQAHVLTGTLTVAVGVLY 527  
 DB 301 AVYDCLALSPAGPSDLTATRLGCAHFSTLPASPLCHALQAHVLTGTLTVAVGVLY 360  
 QY 528 AALVETVALVRGAGNGRPLKLSHVQSQTNCGSPSTPKAHPRRSPRRPQSCSID 587  
 DB 361 AALVETVALVRGAGNGRPLKLSHVQSQTNCGSPSTPKAHPRRSPRRPQSCSID 420  
 QY 588 LGDACCYGYARRLGAMARRSHSVHGLLGACRCGVGSAEPLRESVV 635  
 DB 421 LGDACCYGYARRLGAMARRSHSVHGLLGACRCGVGSAEPLRESVV 468

RESULT 12  
 ADL78778  
 ID ADL78778 standard; protein; 468 AA.  
 AC ADL78778;  
 XX 20-MAY-2004 (first entry)  
 DE Albumin fusion protein related therapeutic protein X, SEQ ID NO 2260.  
 XX albumin fusion protein; cyostatic; antinaemic; antiarthritic;

KW antiasthmatic; anti-HIV; immunosuppressive; antiinflammatory;  
 KW antipsoriatic; antibacterial; osteopathic; dermatological; antigout;  
 KW immunomodulator; antiarrhythmic; cardiac; neurologic; antiplatelet;  
 KW nephrotoxic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;  
 KW antidiabetic; anabolic; hypertensive; vulnery; gene therapy; cancer;  
 KW reproductive system disorder; therapeutic protein.  
 XX Unidentified.  
 OS US2004010134-A1.  
 XX  
 PD 15-JAN-2004.  
 XX  
 PF 12-APR-2001; 2001US-00833245.  
 XX  
 PR 12-APR-2000; 2000US-0229358P.  
 PR 25-APR-2000; 2000US-0195384P.  
 PR 21-DEC-2000; 2000US-0256931P.  
 XX  
 PA (ROSE/) ROSEN C A.  
 PA (HASE/) HASELTINE W A.  
 XX  
 PI Rosen CA, Haseltine WA;  
 XX WPI; 2004-090519/09.  
 DR  
 XX New albumin fusion proteins, useful for diagnosing, treating, preventing  
 PT or ameliorating diseases or disorders e.g. cancer, anemia, arthritis,  
 PT asthma, inflammatory bowel disease or Alzheimer's disease.  
 XX  
 PS Disclosure; SEQ ID NO 2260; 279pp; English.

CC The invention relates to a novel albumin fusion protein. The invention  
 CC further relates to: a composition comprising the albumin fusion protein  
 CC and a pharmaceutical carrier; a kit comprising the albumin fusion protein  
 CC and albumin fusion protein formula; a method of treating a disease or  
 CC disorder in a patient comprising the step of administering the albumin  
 CC fusion protein; a method of treating a patient with a disease or disorder  
 CC that is modulated by Therapeutic protein X, or its fragment or variant;  
 CC a method of extending the shelf life of Therapeutic protein X, or its  
 CC fragment or variant; a nucleic acid molecule comprising a polynucleotide  
 CC sequence encoding the albumin fusion protein; a vector comprising the  
 CC nucleic acid molecule of the albumin fusion protein; and a host cell  
 CC comprising the nucleic acid molecule of the albumin fusion protein. The  
 CC albumin fusion protein and its compositions have the following  
 CC activities: cyostatic, antinaemic, antiarthritic, antiasthmatic, anti-  
 CC HIV, immunosuppressive, antiinflammatory, antipsoriatic, antibacterial,  
 CC osteopathic, dermatological, antigout, immunomodulator, antiarrhythmic,  
 CC cardiac, neurologic, antiplatelet, nephrotoxic, uropathic,  
 CC neuroprotective, antiparkinsonian, tranquilizer, antidiabetic, anabolic,  
 CC hypertensive, and vulnery. The albumin fusion protein nucleic acid may  
 CC be used in gene therapy to treat disorders. The albumin fusion protein is  
 CC useful for diagnosing, treating, preventing or ameliorating diseases or  
 CC disorders comprising indication: Y. The diseases or disorders include:  
 CC cancer (e.g. leukemia, colon, bone, breast, liver or lung cancer),  
 CC immune or haematopoietic diseases (e.g. anemia, Hodgkin's disease, acute  
 CC lymphocytic anaemia, multiple myeloma, arthritis, asthma, AIDS,  
 CC autoimmune disease, inflammatory bowel disease, psoriasis or Lyme  
 CC disease), reproductive system disorders (e.g. prostaticitis, inguinal  
 CC hernia, varicocele, penile carcinoma, ovarian adenocarcinoma or Sertoli-  
 CC Leydig tumours), musculoskeletal diseases (e.g. giant cell tumours,  
 CC Paget's disease, systemic lupus erythematosus, gout, muscular dystrophy  
 CC or cachexia), cardiovascular disease (e.g. thrombomas, heart disease,  
 CC arrhythmia, cardiac arrest, heat valve disease, hypernatraemia or  
 CC hyponatraemia), mixed foetal diseases (e.g. foetal alcohol syndrome,  
 CC Down's syndrome, Patau syndrome, Turner's syndrome, Apert syndrome or Tay-  
 CC Sachs disease), excretory diseases (e.g. urinary incontinence, urinary  
 CC tract infections or renal disorders), neural or sensory disease (e.g.  
 CC Alzheimer's disease, Parkinson's disease, cerebral malaria, meningitis,  
 CC cerebellar ataxia, attention deficit disorder, autism or obsessive  
 CC compulsive disorder), respiratory disease (e.g. emphysema, lung cancer or  
 CC occupational lung disease), endocrine diseases (e.g. diabetes, Addison's  
 CC disease or glomerulonephritis), digestive diseases (e.g. portal

CC hypertension, irritable bowel disease, gastric atrophy or pancreatitis)  
 CC or connective tissue or epithelial diseases (e.g. Crohn's disease,  
 CC scleroderma, wound healing or epidermolysis bullosa). This sequence  
 CC represents a therapeutic protein X relating to the albumin fusion protein  
 CC of the invention. The sequence listing data for this specification was  
 CC downloaded from the USPTO website.

XX Sequence 468 AA:

Query Match 74.2%; Score 2475; DB 8; Length 468;  
 Best Local Similarity 99.8%; Pred. No. 6,3e-184;  
 Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 168 MPRLHTLNDHNLIDALPPGAFQQLGSLDLSNRLATLADPLPSRGDAEASAPL 227  
 Db 1 MPRLHTLNDHNLIDALPPGAFQQLGSLDLSNRLATLADPLPSRGDAEASAPL 60  
 QY 228 VLSFSGNPLHNCNCELMLRLARPPDLETGASPPGLAGRYFMAVPEGEFCEPPLIARHT 287  
 Db 61 VLSFSGNPLHNCNCELMLRLARPPDLETGASPPGLAGRYFMAVPEGEFCEPPLIARHT 120  
 QY 288 QRLWVEGGQATLRRCRALGDPATMHWVGPDDRLVGNSSRAAPFNGTLEIGVTGAGDAG 347  
 Db 121 QRLWVEGGQATLRRCRALGDPATMHWVGPDDRLVGNSSRAAPFNGTLEIGATGAGDAG 180  
 QY 348 GYTCTATNPGEATARVELRVALPHGNSAAGRGPSDIAASATAEGGTLESEP 407  
 Db 181 GYTCTATNPGEATARVELRVALPHGNSAAGRGPSDIAASATAEGGTLESEP 240  
 QY 408 AVQVTEVTATSGLVSWGPGRPADPVMFQIQVNSSEDETLIYRIVPASSHHFLKHLVPG 467  
 Db 241 AVQVTEVTATSGLVSWGPGRPADPVMFQIQVNSSEDETLIYRIVPASSHHFLKHLVPG 300  
 QY 468 ADVDLCLLASPAAGPSDLTATRLGCAHPSSTLPASPLCHALQAHVLGTLTVAVGGLV 527  
 Db 301 ADVDLCLLASPAAGPSDLTATRLGCAHPSSTLPASPLCHALQAHVLGTLTVAVGGLV 360  
 QY 528 AALLVFTVALIVRGKGNRRLPKLSHNSQTMGSPPTPKAHPSPSPRRPORSGLD 587  
 Db 361 AALLVFTVALIVRGKGNRRLPKLSHNSQTMGSPPTPKAHPSPSPRRPORSGLD 420  
 QY 588 LGDAGCGYARLRGAMARRSHSVHGGILGACGKGVGSAERLEESYV 635  
 Db 421 LGDAGCGYARLRGAMARRSHSVHGGILGACGKGVGSAERLEESYV 468

RESULT 13  
 AAB12448  
 ID AAB12448 standard; protein; 785 AA.

XX AC AAB12448;  
 XX 19-OCT-2000 (first entry)  
 XX DE Human hh00149 protein SEQ ID NO:4.  
 XX KW Human; ubiquitin-like protein; 149Y2H#151; hh00149; brain;  
 XX KM 2-hybrid screening; neuroprotective; signal transducer;  
 XX nervous system disease; diagnosis.

OS Homo sapiens.  
 XX XX  
 XX PN W0200031255-A1.  
 XX XX  
 XX PD 02-JUN-2000.  
 XX XX  
 XX PF 18-NOV-1999; 99WO-JP006448.  
 XX XX  
 XX PR 20-NOV-1998; 98JP-00331701.  
 XX XX  
 XX PA \*(CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.  
 XX XX  
 XX PI Funahashi S, Miyata S;

XX WPI, 2000-400066/34.  
 DR N-PSDB; AAA60605.  
 XX XX  
 PT Gene encoding ubiquitin-like protein which interacts with protein  
 PT hh00149, useful in the diagnosis and treatment of diseases associated  
 PT with the nervous system.

XX Example 2; Page 74-79; 88pp; Japanese.

PS The present invention describes a ubiquitin-like protein, designated  
 CC 149Y2H#151, which interacts with protein hh00149 expressed specifically  
 CC in the brain. The 149Y2H#151 protein has neuroprotective activity, and is  
 CC a signal transducer. The 149Y2H#151 gene and encoded protein are useful  
 CC in the diagnosis and treatment of diseases associated with the nervous  
 CC system. The protein can interact with protein hh00149 expressed  
 CC specifically in the brain. The present sequence represents the human  
 CC hh00149 protein

XX Sequence 785 AA:

Query Match 49.8%; Score 1662.5; DB 3; Length 785;  
 Best Local Similarity 46.4%; Pred. No. 2.5e-120;  
 Matches 364; Conservative 83; Mismatches 178; Indels 159; Gaps 13;

QY 9 LLAAGAA-----ACPLPCVCONLSESLSTLCARGLLPFPNVDRRTVELADNFIQAL 63  
 Db 4 LLAAGAAFAVADACPCYTCVCONLSESLSTLCARGLLPFPNVDRRTVELADNFIQAL 63  
 QY 64 GPPDFPNNMTGLVDLTLSRNAITRIGARFGDLESLSLHLDGNRLVELGTSLRGPNVQL 123  
 Db 64 SRQDFPNNMTGLVDLTLSRNTISHIQPFSLDLSRLSHDNRPLSLGDTLRGAVNLQ 123  
 QY 124 HLLSGNQLGRIAPGAFDFLSELDLDSYNNLRQVPMAGIGAMALHTLNDHNLIDA 183  
 Db 124 HLLVNNNOGLGIDAEAFEDFLTELDLDSYNNLHGLPMDSVRRMNLHQLSDHNLIDH 183  
 QY 184 LPFGAFQQLGSLDLSNRLATLADPLPSRGDAEASAPFNGTLEIGVTGAGDAG 239  
 Db 184 IAGTRADLOKLARLDITSNRLQKLPDPPIFASQASALUTAPFAPLPSFGGNPLHNCN 243  
 QY 240 CELLMRLRLARPPDLETGASPPGLAGRYFMAVPEGEFCEPPLIARHTQRLWVEGGQAT 299  
 Db 244 CELLMRLRLERDDLETGASPPGLAGRYFMAVPEGEFCEPPLIARHTQRLWVEGGQAT 303  
 QY 300 LRCRALGDPATMHWVGPDDRLVGNSSRAAPFNGTLEIGVTGAGDAGYTCIATNPAGE 359  
 Db 304 LKCKAIGDPSPLIHWAAPDDRILVGNSSRTAVYDNGTIDITTTSDSGAFTCIANAAGE 363  
 QY 360 ATARVELRVALPHGNSAAGRGPGP--SDIAASATAEGGTLESEP-----AVQ 410  
 Db 364 ATAMVERISIVQLHLSNSTRTAPPKRRLSDITGSSKTSRGSGSGGGEPPKSPERAVL 423  
 QY 411 VTEVTATSGLVSWGPGRPADPVMFQIQVNSSEDETLIYRIVPASSHHFLKHLVPGADY 470  
 Db 424 VSEVTTTSAIVKWSVSKGAPRVKMYQLQYNCSDDEVILYIMIPASNKAFFVNNNVSGGY 483  
 QY 471 DLCLLASPAAGPSDLTATRLGCAHPSSTLPASPLCHALQAHVLGTLTVAVGGLVVAL 530  
 Db 484 DLCLLAMMDDTA--TTLTATNIVGCAOPTYADYFQCSMSQILGGTILVIGGIYATL 542  
 QY 531 LVFTVALIVRGKGNRRLPKL-----SHVOSQTMG-----PSPTPRAHPPRSP----- 577  
 Db 543 LVFTVALIVRGKGNRRLPKL-----SHVOSQTMG-----PSPTPRAHPPRSP----- 577  
 QY 578 -----PR----- 580  
 Db 602 NELLDFTASLARASDSSSSSLGSGEAGLGRAFWIRIPSAFPRKPSLDRLMGAFASIDL 661  
 QY 581 -----ORSCSLD 587  
 Db 662 KSGRKEBLDSRTPAGGACTSARGHHSREPLIGPPAARASLLPLPLBGAKAKSHSFD 721

```

QY 588 LGD-----AGCYGARLGGAMARRSHVHGGLG-----AGCRGVGSARLE 631
DB 722 MGFFAAAGVAVGGYSPPRKVSINWTKRSLSYNGWLLPFESDLVGAGGTGSSSEWVM 781
QY 632 ESVV 635
DB 782 ESTV 785

RESULT 14
AAB09968
ID AAB09968 standard; protein; 789 AA.
AC AAB09968;
XX 19-OCT-2000 (first entry)
DE Human brain-specific transmembrane glycoprotein.
XX Transmembrane glycoprotein; human; brain; cellular signal transducer;
XX neuroactive; neuroprotective; cerebroprotective; drug development;
XX treatment; nervous disease; diagnostic.
OS Homo sapiens.
XX WO200031256-A1.
XX 02-JUN-2000.
XX 18-NOV-1999; 99WO-JP006449.
XX 20-NOV-1998; 98JP-00331727.
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX (KAZU-) KAZUSA DNA RES INST.
XX Funahashi S, Miyata S, Nomura N, Nagase T, Ohara O;
XX MPI; 2000-411954/35.
XX N-PSDB; AAA40083.
XX PT Gene encoding a brain-specific transmembrane glycoprotein that has a
XX typical PDZ protein binding motif and functions as a cellular signal
XX transducer, useful in developing drugs for treating nervous diseases.
XX Claim 1a; Page 52-57; 63pp; Japanese.
XX CC This invention describes a novel gene encoding a human brain-specific
XX transmembrane glycoprotein that has a typical PDZ protein binding motif
XX and functions as a cellular signal transducer. The product of the
XX invention has neuroactive, neuroprotective and cerebroprotective
XX activity. The gene and encoded protein are useful in developing drugs for
XX treating nervous diseases, and also for studying functions of the nervous
XX system or onset mechanism of nerve-related diseases. They may also be
XX used for the production of diagnostic reagents such as oligonucleotide
XX probes and antibodies for detecting proteins comprising PDZ domains and
XX the nucleic acids that encode them. They may also be used to produce and
XX identify modulators of the proteins expression and activity such as
XX antisense sequences and antibodies. This sequence represents the human
XX brain-specific transmembrane glycoprotein described in the method of the
XX invention
XX
XX Sequence 789 AA;
SQ

```

```

Query Match 49.8%; Score 1662.5; DB 3; Length 789;
Best Local Similarity 46.4%; Pred. No.2.5e-120;
Matches 364; Conservative 83; Mismatches 178; Indels 159; Gaps 13;

```

```

QY 9 LTAGSAA-----ACPLPCVCOMLSBSISLTCARHGLLFVPPNVDRRTVELRLDNFIQAL 63
DB 8 LIAFGMAFAVVDACPKVCYCCONLSBSLGLTCCPSKGLLFPPDIDRTVELRLGNFIHII 67
QY 64 GPPDFNNMTGLVDLTLNRNATITIGARAFGDLESLSLHLDGNRLVELGTGSLRGPVNIQ 123

```

```

DB 68 SRQDFANMTGLVDLTLNRNATITISHIQFSPFLDELSLSLHLDNRRLDSLGGDTLRGLVNIQ 127
QY 124 HLISGNOIGRIAPGAFDFPLESLEDLDDSYNNLRQVPMAIGAMPALHTLINDHNIIDA 163
DB 128 HLIVNNNOIGGIADAEFAFEDFLTLLEDLDSYNNLHGLPMDSVRRMNLHQLSIDHNLDH 187
QY 184 LPPGAFAOIGOLSRDLDTSNRLATTAPDPLFSGRDA--EASP--APVLVSFGNPLHCN 239
DB 188 IABGTADLOKLARLDLTLNRLQKLPDPPLFASQASALTATFPAPPLSFGGNPLHCN 247
QY 240 CELLMRLRLARPDDLETCASPPGLAGRYFWAVEPECFSECPPLIARHTORLWLLEGORAT 299
DB 248 CELLMRLRLERDDLETCSPGGLKGRYFMHWRREEFVECPPLITHTIKLVLBEQAAAT 307
QY 300 LRCRALGDPAFTMEWVGPDRLVGNSSRRAPAPNGTLEIGVTAGDAGGYCTIATNPAGE 359
DB 308 LKCKAIGDPSPLIHWVAPDDRDLVGNSSRTAVYNGTLIDIFITTSQDSGAFCTAANAAGE 367
QY 360 ATARVELRLVALPHGNSSAEGRRPGR--SDIAASARTAAEGTLESEPR-----AYQ 410
DB 368 ATAMVEVSIQVLPILSNSTSRTPAPKRSLSDLTGSSKTSRGGSGGGEPPKSPRRAVL 427
QY 411 VTEVTATSGLVSKGPRPADPYWMPQIQYNSEDEFTLIYRIYAPASSHHEFLKLHVPQADY 470
DB 428 VSEVTTSALVKSVSASKAPRYVMYQIQNCSDDEVLIYRMIPASKAKAFVNNLVSGTGY 487
QY 471 DICTLALSPAAGSDTLATRLIGCAHFTLPSPLCHALQAHVLTGTLVAVGCVLVAL 530
DB 488 DICVLAMWDDTA--TTLTATNIVGCAQFFRKADYPOQCSHNSQLIGTMTLIVIGIIVATL 546
QY 531 LVFTVALLVGRGAGNGRLPLKL---SHVSGOTNG---PSPTKAPRRSP-----577
DB 547 LVFTVILMRYK--VCNHEAPSKMAAIVSNVSGTNGAOPPPPSAPAGAPQGPVKVVR 605
QY 578 -----PRP-----580
DB 606 NELLDFTLASLRASDSSSSSSLSGSGEAGLGRAPWRIPPSAPRPKSLRLMGAFAFLDL 665
QY 581 -----ORCSLSD 587
DB 666 KQQRKEELDSRTPAGRGAGTSARGHSDREPLGPAPARASLLPLPLEGKAKRSHSPD 725
QY 588 LGD-----AGCYGARLGGAMARRSHVHGGLG-----AGCRGVGSARLE 631
DB 726 MGFFAAAGVAVPGYSPPRKVSINWTKRSLSYNGWLLPFESDLVGAGGTGSSSEWVM 785
QY 632 ESVV 635
DB 786 ESTV 789

RESULT 15
ADF69106
ID ADF69106 standard; protein; 789 AA.
XX ADF69106;
AC ADF69106;
XX 12-FEB-2004 (first entry)
DE Human MP53 protein sequence SEQ ID NO:76.
XX
XX p53 pathway modulating agent; MP53; p53 modulator; cytostatic;
XX gene therapy; cancer; human.
XX Homo sapiens.
XX OS Homo sapiens.
XX PN WO2003083047-A2.
XX PD 09-OCT-2003.
XX 28-FEB-2003; 2003WO-US006025.
XX

```



PR 01-MAR-2002; 2002US-0361196P.  
 XX (EXEL-) EXELIXIS INC.  
 XX Belvin M, Francis-Lang H, Friedman L, Plowman GD, Heuer TS, Li D;  
 PI Funke RP;  
 XX MPI: 2003-812540/76.  
 DR N-PSDB; ADF69162.  
 XX  
 PT Identifying a candidate p53 pathway modulating agent for treating e.g.,  
 PT cancer by contacting an assay system comprising a Mp53 polypeptide or  
 PT nucleic acid with a test agent and detecting a test agent-biased  
 PT activity.  
 PS Example; SEQ ID NO 76; 406pp; English.  
 XX  
 CC The present invention describes a method for identifying a candidate p53  
 CC pathway modulating agent, which comprises: (a) providing an assay system  
 CC comprising a Mp53 (modulator of p53) polypeptide or nucleic acid or its  
 CC fragment or derivative; (b) contacting the assay system with a test agent  
 CC under conditions where the system provides a reference activity except in  
 CC the presence of the test agent; and (c) detecting a test agent-biased  
 CC activity, where a difference between the test agent-biased activity and  
 CC the reference activity identifies the test agent as a candidate p53  
 CC pathway modulating agent. Also described: (1) modulating the p53 pathway  
 CC of a cell; (2) modulating the p53 pathway in a mammalian cell; and (3)  
 CC diagnosing a disease in a patient. Mp53 has cytosolic activity, and can  
 CC be used in gene therapy. The method is useful for identifying a candidate  
 CC p53 pathway modulating agent for preparing a composition for diagnosing  
 CC or treating e.g., cancer. The present sequence represents a human Mp53  
 CC protein, which is used in the exemplification of the present invention.  
 XX  
 SQ Sequence 789 AA;

Query Match 49.8%; Score 1662.5; DB 7; Length 789;  
 Best Local Similarity 46.4%; Pred. No. 2.5e-120;  
 Matches 364; Conservative 83; Mismatches 178; Indels 159; Gaps 13;

QY 9 LLASGAA-----ACPLECVCONLSESLTLCANRGLFVPPNVDRRTVELRIADNFIQAL 63  
 DB 8 LLAFGMAFAVAVDACPKYCVCONLSESLGTLCPKGLFVPPDIDRTVELRLGDNFIHI 67  
 QY 64 GPPDFRMTGTVDLTLSRNATIRIGARAFGDLRSIHLNRLVEIGTSLRGPNVLO 123  
 DB 68 SRQDFAMTGTVDLTLSRNTSHIQPSFLDLSRLSLHDSNRPLSGEPTLRGLVNLQ 127  
 QY 124 HLISGQLGRIAGAFDPLESLDLISYNNLRQVPMAGIGAMPALHTLNDNLIDA 183  
 DB 128 HLIYNNNQGGIADAEFEDFLTELDLSYNNLHGLPMDSVRMVNLHQSLDNLIDH 187  
 QY 184 LPPGAFQQLQGLSLDLSNRNLATLAPDPLSRGRDA--EASP--APLVLSFGSNPLHCN 239  
 DB 188 IAGTFADLOGLARLDLSNRLOKLPDPIFARSQASALTATPPAPPLSFGSNPLHCN 247  
 QY 240 CELLMRLRLARPDDLETCAAPGLAGRYFMVAPGEFSCPEPLIARTORLWLVGQORAT 299  
 DB 248 CELLMRLRLERDDLETGSGPGKGRYFMVAREEFVCEPPLITQHTHKLIVLEGQAT 307  
 QY 300 LRCFALDPAFTMTWVGDBDLVGNSSRARAFPNGTLEIGVTAGDAGYTCIATNPAGE 359  
 DB 308 LKCAVIGDPPSLIHMVAPDDLVGNSSRTAVYDNGTLDIFITISQDSGAFCIANNAAGE 367  
 QY 360 ATAVEELRLALPHGGSNAAGCRPGP--SDIASARTAAEGEGLSESP-----AVQ 410  
 DB 368 ATANVEVSIQVLPPLSSTSRKAPKSLSDITGSSSKTSRGSGSGGGEPPKSPPERAVL 427  
 QY 411 VTEVATSGLVSWGPGPADVMMFOIQYNSSEDETLIYRIVPASSHFLKHLVPGADY 470  
 DB 428 VSEVITTSALYKMSVKSAPRYVMYQIQYNSDDEVLIRMIIPASNKAFAVNNLVSGTGY 487  
 QY 471 DLCLALSPAAGPSDLTATRLIGCAHRSSTLPASPLCHALQAHVIGTLVAVGVVAAL 530

DB 488 DLCLAMMDPTA--TTLTATNIVGCAQFTKADYPPQCSMHSQILGTMILVIGIIVATL 546  
 QY 531 LVFTVALIVRGAGNGRLPLKL-----SHVQSQTNGA---PPTPKAHPRSP----- 577  
 DB 547 LVFTVILIMVRYK-VCHNEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPGPKVVR 605  
 QY 578 -----PP----- 580  
 DB 606 NEILDFTASLARSDSSSSSLGSGEAGLGRAPWRIPPGAPRPKPSLDRLMGAFASLDL 665  
 QY 581 -----ORSCSLD 587  
 DB 666 KSQRKEILDSTRTPAGRGACTSARGHHSDBREPLLPAPARASLLPLBGAKRSHSFD 725  
 QY 588 LGD-----AGCYGARPLGAMARSHSVHGLG-----AGCRGVGSARLE 631  
 DB 726 MGFMAAAAGVVPVGGYSPPRKXSNITKRSLSVNGMLPFESDVLVGARGTSGSSSEWM 785  
 QY 632 ESVV 635  
 DB 786 ESTV 789

Search completed: November 17, 2004, 22:36:45  
 Job time : 77 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 22:40:15 ; Search time 141 Seconds

(Without alignments)  
1594.830 Million cell updates/sec

Title: US-10-071-879-10

Perfect score: 3336

Sequence: 1 MAPPLLLLLIAGAAACPLP.....LGAQCRGVGSARLEESVY 635

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1570615 segs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/prodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/prodata/2/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/prodata/2/pubpaa/US11\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/prodata/2/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3319	96.5	778	14	US-10-093-463-192
2	3026.5	90.7	647	14	US-10-093-463-188
3	2965	89.9	565	9	US-09-768-826-58
4	2773.5	83.1	566	14	US-10-093-463-190
5	2773	83.1	551	9	US-09-897-214-8
6	2475	74.2	468	9	US-09-768-826-40
7	2475	74.2	468	11	US-09-893-245-2260
8	1662.5	49.8	832	10	US-09-732-436-24
9	1662.5	49.8	832	14	US-10-004-378A-41
10	1657.5	49.7	789	14	US-10-004-378A-42
11	1655.5	49.6	789	14	US-10-291-172-261
12	1655.5	49.6	789	14	US-10-221-278-261
13	1630	48.9	788	14	US-10-004-378A-43

14	1557	46.7	628	9	US-09-815-626-2
15	1557	46.7	628	14	US-10-145-586-61
16	1556	46.6	628	10	US-09-732-436-6
17	1556	46.6	628	14	US-10-221-097-45
18	1546.5	46.4	627	14	US-10-245-752-100
19	1546.5	46.4	627	14	US-10-245-859-100
20	1546.5	46.4	627	14	US-10-245-103-100
21	1546.5	46.4	627	14	US-10-245-107-100
22	1546.5	46.4	627	14	US-10-245-143-100
23	1546.5	46.4	627	14	US-10-245-771-100
24	1546.5	46.4	627	14	US-10-245-851-100
25	1546.5	46.4	627	14	US-10-245-883-100
26	1546.5	46.4	627	14	US-10-237-535-100
27	1546.5	46.4	627	14	US-10-238-183-100
28	1546.5	46.4	627	14	US-10-238-283-100
29	1546.5	46.4	627	14	US-10-238-370-100
30	1546.5	46.4	627	14	US-10-245-147-100
31	1546.5	46.4	627	14	US-10-245-109-100
32	1546.5	46.4	627	14	US-10-245-730-100
33	1546.5	46.4	627	14	US-10-245-739-100
34	1546.5	46.4	627	14	US-10-245-210-100
35	1546.5	46.4	627	14	US-10-235-156-100
36	1546.5	46.4	627	14	US-10-243-024-100
37	1546.5	46.4	627	14	US-10-243-409-100
38	1546.5	46.4	627	14	US-10-245-621-100
39	1546.5	46.4	627	14	US-10-245-880-100
40	1546.5	46.4	627	14	US-10-245-033-100
41	1546.5	46.4	627	14	US-10-243-095-100
42	1546.5	46.4	627	14	US-10-245-185-100
43	1546.5	46.4	627	14	US-10-245-427-100
44	1546.5	46.4	627	14	US-10-245-473-100
45	1546.5	46.4	627	14	US-10-245-770-100

#### ALIGNMENTS

RESULT 1  
US-10-093-463-192  
Sequence 192, Application US/10093463  
Publication No. US20030208039A1  
GENERAL INFORMATION:  
APPLICANT: Padigar, Muralidhara  
APPLICANT: Shenoy, Suresh  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Gusev, Vladimir  
APPLICANT: Pochart, Pascal  
APPLICANT: Zhong, Mei  
APPLICANT: Rastelli, Luca  
APPLICANT: Mezes, Peter  
APPLICANT: Smithson, Glenda  
APPLICANT: Guo, Xiaojia  
APPLICANT: Gerlach, Valerie  
APPLICANT: Casman, Stacie  
APPLICANT: Boldog, Ferenc  
APPLICANT: Li, Li  
APPLICANT: Zerhusen, Bryan  
APPLICANT: Tchiernev, Velizar  
APPLICANT: Gangoli, Esna  
APPLICANT: Vernet, Corine  
APPLICANT: Pena, Carol  
APPLICANT: Burgess, Catherine  
APPLICANT: Liu, Xiaohong  
APPLICANT: Spytek, Kimberly  
APPLICANT: Gorman, Linda  
APPLICANT: Spaderna, Steven  
APPLICANT: Voss, Edward  
APPLICANT: Malyankar, Uriel  
APPLICANT: Anderson, David  
APPLICANT: Patturajan, Meera  
APPLICANT: Miller, Charles  
APPLICANT: Taupier, Raymond J. Jr.  
TITLE OF INVENTION: No. US20030208039A1 Antibodies that Bind to Antigenic Polypeptide

TITLE OF INVENTION: Encoding The Antigens, and Methods of Use.  
FILE REFERENCE: 21402-290A (Cura 590AT)  
CURRENT APPLICATION NUMBER: US/10/093,463  
CURRENT FILING DATE: 2002-06-24  
PRIOR APPLICATION NUMBER: 60/283,675  
PRIOR FILING DATE: 2001-04-14  
PRIOR APPLICATION NUMBER: 60/338,092  
PRIOR FILING DATE: 2001-12-03  
PRIOR APPLICATION NUMBER: 60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/274,101  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/325,681  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: 60/304,354  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/279,995  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: 60/294,899  
PRIOR FILING DATE: 2001-05-31  
PRIOR APPLICATION NUMBER: 60/287,424  
PRIOR FILING DATE: 2001-04-30  
PRIOR APPLICATION NUMBER: 60/299,027  
PRIOR FILING DATE: 2001-06-18  
PRIOR APPLICATION NUMBER: 60/309,198  
PRIOR FILING DATE: 2001-07-31  
PRIOR APPLICATION NUMBER: 60/281,194  
PRIOR FILING DATE: 2001-04-04  
PRIOR APPLICATION NUMBER: 60/274,194  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/274,849  
PRIOR FILING DATE: 2001-03-09  
PRIOR APPLICATION NUMBER: 60/330,380  
PRIOR FILING DATE: 2001-10-18  
PRIOR APPLICATION NUMBER: 60/275,235  
PRIOR FILING DATE: 2001-03-12  
PRIOR APPLICATION NUMBER: 60/288,342  
PRIOR FILING DATE: 2001-05-03  
PRIOR APPLICATION NUMBER: 60/275,578  
PRIOR FILING DATE: 2001-03-13  
NUMBER OF SEQ ID NOS: 370  
SOFTWARE: Patentm Ver. 2.1  
SEQ ID NO 192  
LENGTH: 778  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-093-463-192

Query Match 96.5%; Score 3219; DB 14; Length 778;  
Best Local Similarity 96.1%; Pred. No. 1.9e-224;  
Matches 617; Conservative 2; Mismatches 5; Indels 18; Gaps 1;

QY 1 MAPPLILLILASGAACPLPCVCNLTSESLTLCARGLLFVPPNVDRTVEIRLADNFI 60  
DB 1 MAPPLILLILASGAACPLPCVCNLTSESLTLCARGLLFVPPNVDRTVEIRLADNFI 60  
QY 61 QALGPFRFMTGTVDLTLSNATTRIGARFGLESIRSHLDGNLVEIGTSLGPV 120  
DB 61 QALGPFRFMTGTVDLTLSNATTRIGARFGLESIRSHLDGNLVEIGTSLGPV 120  
QY 121 NLQHLIISGNQIGRIAPGAPDFLESLEDLDSYNNLRQVWAGIGAMPALHTLNDHL 180  
DB 121 NLQHLIISGNQIGRIAPGAPDFLESLEDLDSYNNLRQVWAGIGAMPALHTLNDHL 180  
QY 181 IDALPPGAFAQLQGLSRDLTSSNRIATLAPDPLFSRGDAASPAIVLSFGNPLHCNC 240  
DB 181 IDALPPGAFAQLQGLSRDLTSSNRIATLAPDPLFSRGDAASPAIVLSFGNPLHCNC 240  
QY 241 ELLMLRLARPDDETACSPGGLAGRYWAVPGEFSGEPPLIARHQRMLVLEGGQATL 300  
DB 241 ELLMLRLARPDDETACSPGGLAGRYWAVPGEFSGEPPLIARHQRMLVLEGGQATL 300  
QY 301 RCRLGDPAPFTMHVGPDDRIVGNSRRARFPNGLTETGVTGAGDAGGYTCIATNPAGEA 360

DB 301 RCRLGDPAPFTMHVGPDDRIVGNSRRARFPNGLTETGVTGAGDAGGYTCIATNPAGEA 360  
QY 361 TARVELRLVALPHGNNSSAEGR-----PGSDIASATTAEGEET 402  
DB 361 TARVELRLVALPHGNNSSAEGRPGPRTSAPWMEODCRGPGPSDIAASATTAEGEET 420  
QY 403 LSEPAVQVTEVTATSGVSWGPRPADVWMEFOIYNNSEDETILYRIVPASSHFLK 462  
DB 403 LSEPAVQVTEVTATSGVSWGPRPADVWMEFOIYNNSEDETILYRIVPASSHFLK 480  
QY 463 HLPAGADYDLCILALSPAAGPSDLTATRLGCAHFTTLPASPLCHALQAHVIGTITVAV 522  
DB 463 HLPAGADYDLCILALSPAAGPSDLTATRLGCAHFTTLPASPLCHALQAHVIGTITVAV 540  
QY 523 GGTVVALLVFTVALLVRRGANGRLPLKSHVOSQTNMGSSPTPKAHPPSPRPOR 582  
DB 523 GGTVVALLVFTVALLVRRGANGRLPLKSHVOSQTNMGSSPTPKAHPPSPRPOR 600  
QY 583 SCSLDLGDAGCYGARRLGAMARRSHSVHGGLGAGCRGVG 624  
DB 583 SCSLDLGDAGCYGARRLGAMARRSHSVHGGLGAGCRGVG 642

RESULT 2  
US-10-093-463-188  
Sequence 188, Application US/10093463  
Publication No. US20030208039A1  
GENERAL INFORMATION:  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Shenoy, Suresh  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Gusev, Vladimir  
APPLICANT: Rochart, Pascal  
APPLICANT: Zhong, Mei  
APPLICANT: Rastelli, Luca  
APPLICANT: Mezes, Peter  
APPLICANT: Smithson, Glenda  
APPLICANT: Guo, Xiaojia  
APPLICANT: Gerlach, Valerie  
APPLICANT: Casman, Stacie  
APPLICANT: Boldog, Ferenc  
APPLICANT: Li, Li  
APPLICANT: Zerhusen, Bryan  
APPLICANT: Tchernen, Velizar  
APPLICANT: Gangolli, Resha  
APPLICANT: Verne, Corine  
APPLICANT: Pena, Carol  
APPLICANT: Burgess, Catherine  
APPLICANT: Liu, Xiaohong  
APPLICANT: Spytek, Kimberly  
APPLICANT: Gorman, Linda  
APPLICANT: Spaderna, Steven  
APPLICANT: Voss, Edward  
APPLICANT: Malyankar, Uriel  
APPLICANT: Anderson, David  
APPLICANT: Patuturajan, Meera  
APPLICANT: Miller, Charles  
APPLICANT: Taupier, Raymond J, Jr.  
TITLE OF INVENTION: No. US20030208039A1 Antibodies that Bind to Antigenic Polypept  
FILE REFERENCE: 21402-290A (Cura 590AT)  
CURRENT APPLICATION NUMBER: US/10/093,463  
CURRENT FILING DATE: 2002-06-24  
PRIOR APPLICATION NUMBER: 60/283,675  
PRIOR FILING DATE: 2001-04-14  
PRIOR APPLICATION NUMBER: 60/338,092  
PRIOR FILING DATE: 2001-12-03  
PRIOR APPLICATION NUMBER: 60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/274,101  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/325,681

```

; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/275,578
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 188
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-093-463-188

```

```

Query Match      90.7%; Score 3026.5; DB 14; Length 647;
Best Local Similarity 93.9%; Pred. No. 1.4e-210;
Matches 589; Conservative 0; Mismatches 21; Indels 17; Gaps 3;

```

```

QY 1 MAPPILILLASGAACAPLPCVCQNLSESLTLCARGLTFVPPVNDRTVELRLADFI 60
DB 1 MAPPILILLASGAACAPLPCVCQNLSESLTLCARGLTFVPPVNDRTVELRLADFI 60
QY 61 QALGPDERMTGLVDTLTSRNATRTICARAFGLDESLSLHDGNRLVELGTGSLRGPV 120
DB 61 QALGPDERMTGLVDTLTSRNATRTICARAFGLDESLSLHDGNRLVELGTGSLRGPV 120
QY 121 NLQHLISGNQIGRIAPGAFDDFLESLEDLDSYNNLRQVPWAGIGAMPALHTLNDHNL 180
DB 121 NLQHLISGNQIGRIAPGAFDDFLESLEDLDSYNNLRQVPWAGIGAMPALHTLNDHNL 180
QY 181 IDALPPGAFQIGQLSRDLTNSNRLATLAPDPLFSRGRDAEAPAVLVSFGNPLHCNC 240
DB 181 IDALPPGAFQIGQLSRDLTNSNRLATLAPDPLFSRGRDAEAPAVLVSFGNPLHCNC 240
QY 241 ELWLRLARDDDETCA SPGLAGRYFWAVPEGEFSCPEPLIARHTQRLVLEGGORATL 300
DB 241 ELWLRLARDDDETCA SPGLAGRYFWAVPEGEFSCPEPLIARHTQRLVLEGGORATL 300
QY 301 RCRLADPAPMTMHWGDDRLVGNSSRARAFPNGLTLEIGTGADAGGYTCIATNPAGEA 360
DB 301 RCRLADPAPMTMHWGDDRLVGNSSRARAFPNGLTLEIGTGADAGGYTCIATNPAGEA 360
QY 361 TARVELFVLALPHGNSSAEGRGDPDIAASARTAAEGEGTLESEPAVQVTEVTATSGL 420
DB 361 TARVELFVLALPHGNSSAEGRGDPDIAASARTAAEGEGTLESEPAVQVTEVTATSGL 420
QY 421 VSWGPGPADPVMWFOIQYNSSEDETLIYRIVPASSHFLKHLVPGADVDLCIALASPA 480
DB 421 VSWGPGPADPVMWFOIQYNSSEDETLIYRIVPASSHFLKHLVPGADVDLCIALASPA 480
QY 481 AGPSDLTATRLGCAHFSTLPASPLCHALQAHVGLGTLTVAAGVVAALLVFTVALVLR 540
DB 481 AGPSDLTATRLGCAHFSTLPASPLCHALQAHVGLGTLTVAAGVVAALLVFTVALVLR 540

```

```

DB 481 AGPSDLTATRLGCAHFSTLPASPLCHALQAHVGLGTLTVAAGVVAALLVFTVALVLR 540
QY 541 GRGAGNGRLPLKLSHVQSTNGSPSTPPKAAHPPRSP-----DRPQSCLDYGDAG 592
DB 541 GRGAGNGRLPLKLSHVQSTNGSPSTPPKAAHPPRSPBAAQULSGPCRCVLR----- 596
QY 593 CYGVARRLG-----AMARRSHVHG 614
DB 597 CQAPGRSLQPTTEPLCAKMGAAARGVPEG 623

```

RESULT 3

```

US-09-768-826-58
; Sequence 58, Application US/09768826
; Patent No. US20020012966A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: 18 human secreted proteins
; FILE REFERENCE: P512P1
; CURRENT APPLICATION NUMBER: US/09/768,826
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US00/22350
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/148,759
; PRIOR FILING DATE: 1999-08-16
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO: 58
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (270)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-768-826-58

```

```

Query Match      88.9%; Score 2965; DB 9; Length 565;
Best Local Similarity 99.8%; Pred. No. 3.3e-206;
Matches 564; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 71 MTGLVDTLTSNNAITRICARAFGLDESLSLHDGNRLVELGTGSLRGPVNLQHLISGN 130
DB 1 MTGLVDTLTSNNAITRICARAFGLDESLSLHDGNRLVELGTGSLRGPVNLQHLISGN 60
QY 131 QLGRIAAGAPDDELESLEDLDSYNNLRQVPWAGIGAMPALHTLNDHNLIDALPPGAF 190
DB 131 QLGRIAAGAPDDELESLEDLDSYNNLRQVPWAGIGAMPALHTLNDHNLIDALPPGAF 120
QY 61 QLGRIAAGAPDDELESLEDLDSYNNLRQVPWAGIGAMPALHTLNDHNLIDALPPGAF 120
DB 61 QLGRIAAGAPDDELESLEDLDSYNNLRQVPWAGIGAMPALHTLNDHNLIDALPPGAF 120
QY 191 QLQSLRLDLTNSNRLATLAPDPLFSRGRDAEAPAVLVSFGNPLHCNCGLWLRLAR 250
DB 121 QLQSLRLDLTNSNRLATLAPDPLFSRGRDAEAPAVLVSFGNPLHCNCGLWLRLAR 180
QY 251 PDLLETCA SPGLAGRYFWAVPEGEFSCPEPLIARHTQRLVLEGGORATLRCRALGPAP 310
DB 181 PDLLETCA SPGLAGRYFWAVPEGEFSCPEPLIARHTQRLVLEGGORATLRCRALGPAP 240
QY 311 TMHWGDDRLVGNSSRARAFPNGLTLEIGTGADAGGYTCIATNPAGEATARVELRYLA 370
DB 241 TMHWGDDRLVGNSSRARAFPNGLTLEIGTGADAGGYTCIATNPAGEATARVELRYLA 300
QY 371 LPHGNSSAEGRGDPDIAASARTAAEGEGTLESEPAVQVTEVTATSGVSWGPGPAD 430
DB 301 LPHGNSSAEGRGDPDIAASARTAAEGEGTLESEPAVQVTEVTATSGVSWGPGPAD 360
QY 431 PVMWFOIQYNSSEDETLIYRIVPASSHFLKHLVPGADVDLCIALASPAAGSDTLATR 490
DB 361 PVMWFOIQYNSSEDETLIYRIVPASSHFLKHLVPGADVDLCIALASPAAGSDTLATR 420
QY 491 LIGCAHFSTLPASPLCHALQAHVGLGTLTVAAGVVAALLVFTVALVLRGAGNGRLP 550
DB 421 LIGCAHFSTLPASPLCHALQAHVGLGTLTVAAGVVAALLVFTVALVLRGAGNGRLP 480

```

QY 551 LKLSHVQQTNGGPEPTPKAHPPPRPGSGSLDGDAGCTGYARRLGAMARRSHS 610  
DB 481 LKLSHVQQTNGGPEPTPKAHPPPRPGSGSLDGDAGCTGYARRLGAMARRSHS 540  
QY 611 VHGGILGAGCGGVGSGARRLEESV 635  
DB 541 VHGGILGAGCGGVGSGARRLEESV 565

RESULT 4  
US-10-093-463-190  
Sequence 190, Application US/10093463  
Publication No. US20030208039A1  
GENERAL INFORMATION:  
APPLICANT: Padigar, Muralidhara  
APPLICANT: Shenoy, Suresh  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Gusev, Vladimir  
APPLICANT: Pochart, Pascal  
APPLICANT: Zhong, Mei  
APPLICANT: Rastelli, Luca  
APPLICANT: Mezes, Peter  
APPLICANT: Smithson, Glenda  
APPLICANT: Guo, Xiaojia  
APPLICANT: Gerlach, Valerie  
APPLICANT: Casman, Stacie  
APPLICANT: Boldog, Ferenc  
APPLICANT: Li, Li  
APPLICANT: Zerhusen, Bryan  
APPLICANT: Tchernev, Velizar  
APPLICANT: Gangoli, Esha  
APPLICANT: Vernet, Corine  
APPLICANT: Pena, Carol  
APPLICANT: Burgess, Catherine  
APPLICANT: Liu, Xiaohong  
APPLICANT: Spytek, Kimberly  
APPLICANT: Gormen, Linda  
APPLICANT: Spaderna, Steven  
APPLICANT: Voss, Edward  
APPLICANT: Malyankar, Uriel  
APPLICANT: Anderson, David  
APPLICANT: Patuturajan, Meera  
APPLICANT: Miller, Charles  
APPLICANT: Taupier, Raymond J. Jr.  
TITLE OF INVENTION: No. US20030208039A1 Anticodons that Bind to Antigenic Polypeptide  
TITLE OF INVENTION: Encoding The Antigens, and Methods of Use.  
FILE REFERENCE: 21402-290A (Cura 590AT)  
CURRENT APPLICATION NUMBER: US/10/093,463  
CURRENT FILING DATE: 2002-06-24  
PRIOR APPLICATION NUMBER: 60/283,675  
PRIOR FILING DATE: 2001-04-14  
PRIOR APPLICATION NUMBER: 60/338,092  
PRIOR FILING DATE: 2001-12-03  
PRIOR APPLICATION NUMBER: 60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/274,101  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/325,681  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: 60/304,354  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/279,995  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: 60/294,899  
PRIOR FILING DATE: 2001-05-31  
PRIOR APPLICATION NUMBER: 60/287,424  
PRIOR FILING DATE: 2001-04-30  
PRIOR APPLICATION NUMBER: 60/299,027  
PRIOR FILING DATE: 2001-06-18  
PRIOR APPLICATION NUMBER: 60/309,198  
PRIOR FILING DATE: 2001-07-31  
PRIOR APPLICATION NUMBER: 60/281,194  
PRIOR FILING DATE: 2001-04-04

PRIOR APPLICATION NUMBER: 60/274,194  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/274,849  
PRIOR FILING DATE: 2001-03-09  
PRIOR APPLICATION NUMBER: 60/330,380  
PRIOR FILING DATE: 2001-10-18  
PRIOR APPLICATION NUMBER: 60/275,235  
PRIOR FILING DATE: 2001-03-12  
PRIOR APPLICATION NUMBER: 60/288,342  
PRIOR FILING DATE: 2001-05-03  
PRIOR APPLICATION NUMBER: 60/275,578  
PRIOR FILING DATE: 2001-03-13  
NUMBER OF SEQ ID NOS: 370  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 190  
LENGTH: 566  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-093-463-190

Query Match 83.1%; Score 2773.5; DB 14; Length 566;  
Best Local Similarity 95.4%; Pred. No. 2.5e-192;  
Matches 537; Conservative 4; Mismatches 11; Indels 11; Gaps 2;

QY 1 MAPPLILLASGAACPLPCVCONLSBSLSTLCARHGLFVPPVNDRTVEIRLADNFI 60  
DB 1 MAPPLILLASGAACPLPCVCONLSBSLSTLCARHGLFVPPVNDRTVEIRLADNFI 60  
QY 61 QALGPPDFRMTGLVDLTLSRNAITRIGARAGDLESLSHLDDGRVLVEIGTSIRGPV 120  
DB 61 QALGPPDFRMTGLVDLTLSRNAITRIGARAGDLESLSHLDDGRVLVEIGTSIRGPV 120  
QY 121 NIQHILISGNOLGRIAPGAFDDLESIEDLSYNNLRQVPWAGICAMPALHTLNDHL 180  
DB 121 NIQHILISGNOLGRIAPGAFDDLESIEDLSYNNLRQVPWAGICAMPALHTLNDHL 180  
QY 181 IDALPPGAFQOLGSLRDLTSNRLLATLAPDLSFGKRDAAEPALVLSFGNPLHCNC 240  
DB 181 IDALPPGAFQOLGSLRDLTSNRLLATLAPDLSFGKRDAAEPALVLSFGNPLHCNC 240  
QY 241 ELLMLRLARPDDLETCAEPGLAGRYFPAVVBEGESCEPPLIARTORTLVLEGORATL 300  
DB 241 ELLMLRLARPDDLETCAEPGLAGRYFPAVVBEGESCEPPLIARTORTLVLEGORATL 300  
QY 301 RCRAIGDPAPTMHWGPPDRLVGNSSRARAFPGTLEIGTGAGDAGYTCIATNPAGEA 360  
DB 301 RCRAIGDPAPTMHWGPPDRLVGNSSRARAFPGTLEIGTGAGDAGYTCIATNPAGEA 360  
QY 361 TARVELRLVLPFGNSSAEGRPGPSDIAASRTAAEGEGLSEPPAVQVTEVTATSGL 420  
DB 361 TARVELRLVLPFGNSSAEGRPGPSDIAASRTAAEGEGLSEPPAVQVTEVTATSGL 420  
QY 421 VSWGGRPADPVMWFQIYNSSSEDTLIRIVASSHHFLKLVAGADVDCILALSPA 480  
DB 421 VSWGGRPADPVMWFQIYNSSSEDTLIRIVASSHHFLKLVAGADVDCILALSPA 480  
QY 481 AGPSDLTATRLGCAFPSTLPASPLCHALCAHVLGTLTVAVGVVVAALVFTVALVR 540  
DB 481 AGPSDLTATRLGCAFPSTLPASPLCHALCAHVLGTLTVAVGVVVAALVFTVALVR 540  
QY 541 GRG-----AGNGRLPLKL 553  
DB 540 WRGLPLPWSGGQNRGKSLPLQV 562

RESULT 5  
US-09-897-214-8  
Sequence 8, Application US/09897214  
Patent No. US20020076779A1  
GENERAL INFORMATION:  
APPLICANT: Thayer, Edward C.  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Presnell, Scott R.

; TITLE OF INVENTION: Leucine-Rich Repeat Proteins, Zlrr7,  
 ; FILE REFERENCE: Zlrr8, and Zlrr9  
 ; CURRENT APPLICATION NUMBER: US/09/897,214  
 ; CURRENT FILING DATE: 2001-07-02  
 ; PRIOR APPLICATION NUMBER: 60/215,446  
 ; PRIOR FILING DATE: 2000-06-30  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 8  
 ; LENGTH: 551  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-897-214-8

Query Match 83.1%; Score 2773; DB 9; Length 551;  
 Best Local Similarity 100.0%; Pred. No.2,6e-192;  
 Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MADPDLILLASGAAACPLPCVCONLSESLSTLCARGLLFPPNDRTVELRLADNFI 60  
 DB 1 MADPDLILLASGAAACPLPCVCONLSESLSTLCARGLLFPPNDRTVELRLADNFI 60  
 QY 61 QALGPDEFNMTGLVLTLSRNAITRIGARAFGDLISLSLHLDGNRLVELGTSLRGVY 120  
 DB 61 QALGPDEFNMTGLVLTLSRNAITRIGARAFGDLISLSLHLDGNRLVELGTSLRGVY 120  
 QY 121 NLQHLISGNQLGRIAPGAFDDLESLDLSYNNLRQVPMAGIAMPALHTLNDHL 180  
 DB 121 NLQHLISGNQLGRIAPGAFDDLESLDLSYNNLRQVPMAGIAMPALHTLNDHL 180  
 QY 181 IDALPPGAFQIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVLVSFGSNPLHCNC 240  
 DB 181 IDALPPGAFQIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVLVSFGSNPLHCNC 240  
 QY 241 ELLMLRRLARPDDLETCASPPGIAGRYFWAVPEGEFSCBPPIARHTQRLMWLEGGRAVL 300  
 DB 241 ELLMLRRLARPDDLETCASPPGIAGRYFWAVPEGEFSCBPPIARHTQRLMWLEGGRAVL 300  
 QY 301 RCRALGDPAPTMHWVGPDDRLVGNSSRAAFAPNGTLEIGVAGAGAGVTCIATNPAGEA 360  
 DB 301 RCRALGDPAPTMHWVGPDDRLVGNSSRAAFAPNGTLEIGVAGAGAGVTCIATNPAGEA 360  
 QY 361 TARVELRVLALPHGNSAEGGRPGPSDIAASARTAAEGGTLESEPAVQTEVTATSGT 420  
 DB 361 TARVELRVLALPHGNSAEGGRPGPSDIAASARTAAEGGTLESEPAVQTEVTATSGT 420  
 QY 421 VSMGGRPADPYWMFQIQYNSSEDETLIRIVPASSHHFLKHLVPGADYDCLIALSPA 480  
 DB 421 VSMGGRPADPYWMFQIQYNSSEDETLIRIVPASSHHFLKHLVPGADYDCLIALSPA 480  
 QY 481 AGPSDLTATRLIGCAHFSTLPASPLCHALQAHVLGGTLTVAVGVYVAL 530  
 DB 481 AGPSDLTATRLIGCAHFSTLPASPLCHALQAHVLGGTLTVAVGVYVAL 530

RESULT 6  
 ; US-09-768-826-40  
 ; Sequence 40, Application US/09768826  
 ; Patent No. US20020012966A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shi et al.  
 ; TITLE OF INVENTION: 18 human secreted proteins  
 ; FILE REFERENCE: PF512P1  
 ; CURRENT APPLICATION NUMBER: US/09/768,826  
 ; CURRENT FILING DATE: 2001-01-25  
 ; PRIOR APPLICATION NUMBER: PCT/US00/22350  
 ; PRIOR FILING DATE: 2000-08-15  
 ; PRIOR APPLICATION NUMBER: 60/148,759  
 ; PRIOR FILING DATE: 1999-08-16  
 ; NUMBER OF SEQ ID NOS: 61  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 40

; LENGTH: 468  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-768-826-40

Query Match 74.2%; Score 2475; DB 9; Length 468;  
 Best Local Similarity 99.8%; Pred. No.8,3e-171;  
 Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 168 MPALHTNLIDHNLIDALPPGAFQIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVL 227  
 DB 1 MPALHTNLIDHNLIDALPPGAFQIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVL 227  
 QY 228 VLSFGSNPLHCNCELMLRLARPDDLETCASPPGIAGRYFWAVPEGEFSCBPPIARHT 287  
 DB 228 VLSFGSNPLHCNCELMLRLARPDDLETCASPPGIAGRYFWAVPEGEFSCBPPIARHT 287  
 QY 61 VLSFGSNPLHCNCELMLRLARPDDLETCASPPGIAGRYFWAVPEGEFSCBPPIARHT 120  
 DB 61 VLSFGSNPLHCNCELMLRLARPDDLETCASPPGIAGRYFWAVPEGEFSCBPPIARHT 120  
 QY 288 QRLMWLEGGRAVLRCRAGDPAPTMHWVGPDDRLVGNSSRAAFAPNGTLEIGVAGAG 347  
 DB 288 QRLMWLEGGRAVLRCRAGDPAPTMHWVGPDDRLVGNSSRAAFAPNGTLEIGVAGAG 347  
 QY 121 QRLMWLEGGRAVLRCRAGDPAPTMHWVGPDDRLVGNSSRAAFAPNGTLEIGVAGAG 180  
 DB 121 QRLMWLEGGRAVLRCRAGDPAPTMHWVGPDDRLVGNSSRAAFAPNGTLEIGVAGAG 180  
 QY 348 GTTCIATNPAGEATARVELRVLALPHGNSAEGGRPGPSDIAASARTAAEGGTLESE 407  
 DB 348 GTTCIATNPAGEATARVELRVLALPHGNSAEGGRPGPSDIAASARTAAEGGTLESE 407  
 QY 181 GYTCTIATNPAGEATARVELRVLALPHGNSAEGGRPGPSDIAASARTAAEGGTLESE 240  
 DB 181 GYTCTIATNPAGEATARVELRVLALPHGNSAEGGRPGPSDIAASARTAAEGGTLESE 240  
 QY 408 AVQVTEVTATSGLVSMGGRPADPYWMFQIQYNSSEDETLIRIVPASSHHFLKHLVPG 467  
 DB 408 AVQVTEVTATSGLVSMGGRPADPYWMFQIQYNSSEDETLIRIVPASSHHFLKHLVPG 467  
 QY 241 AVQVTEVTATSGLVSMGGRPADPYWMFQIQYNSSEDETLIRIVPASSHHFLKHLVPG 300  
 DB 241 AVQVTEVTATSGLVSMGGRPADPYWMFQIQYNSSEDETLIRIVPASSHHFLKHLVPG 300  
 QY 468 ADVDCIALLSPPAGPSDLTATRLIGCAHFSTLPASPLCHALQAHVLGGTLTVAVGVY 527  
 DB 468 ADVDCIALLSPPAGPSDLTATRLIGCAHFSTLPASPLCHALQAHVLGGTLTVAVGVY 527  
 QY 301 ADVDCIALLSPPAGPSDLTATRLIGCAHFSTLPASPLCHALQAHVLGGTLTVAVGVY 360  
 DB 301 ADVDCIALLSPPAGPSDLTATRLIGCAHFSTLPASPLCHALQAHVLGGTLTVAVGVY 360  
 QY 528 AALLVFTVALIVRGAGNGRPLPLKLSHVQQTNGGSPPTPAHPPRSPPPRQSCSD 587  
 DB 528 AALLVFTVALIVRGAGNGRPLPLKLSHVQQTNGGSPPTPAHPPRSPPPRQSCSD 587  
 QY 361 AALLVFTVALIVRGAGNGRPLPLKLSHVQQTNGGSPPTPAHPPRSPPPRQSCSD 420  
 DB 361 AALLVFTVALIVRGAGNGRPLPLKLSHVQQTNGGSPPTPAHPPRSPPPRQSCSD 420  
 QY 588 LGDACCYGARPLGAMARSHSVHGLLAGCCRGVGAERLESYV 635  
 DB 588 LGDACCYGARPLGAMARSHSVHGLLAGCCRGVGAERLESYV 635  
 QY 421 LGDACCYGARPLGAMARSHSVHGLLAGCCRGVGAERLESYV 468  
 DB 421 LGDACCYGARPLGAMARSHSVHGLLAGCCRGVGAERLESYV 468

RESULT 7  
 ; US-09-833-245-2260  
 ; Sequence 2260, Application US/09833245  
 ; Publication No. US20040010134A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Human Genome Sciences, Inc.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF546PCT  
 ; CURRENT APPLICATION NUMBER: US/09/833,245  
 ; CURRENT FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229,358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256,931  
 ; PRIOR FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: 60/199,384  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2260  
 ; LENGTH: 468  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-833-245-2260

Query Match 74.2%; Score 2475; DB 11; Length 468;  
 Best Local Similarity 99.8%; Pred. No.8,3e-171;  
 Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 168 MPALHTNLIDHNLIDALPPGAFQIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVL 227  
 DB 1 MPALHTNLIDHNLIDALPPGAFQIGQLSRDLTNSRLATLAPDPLFSRGRDAEASPAVL 227

QY 228 VLSFSGNPLHNCNCELLMRLRLARPDDLETCA SPPLAGRYFMAVPEGEFSCEPPLIART 287  
 Db 61 VLSFSGNPLHNCNCELLMRLRLARPDDLETCA SPPLAGRYFMAVPEGEFSCEPPLIART 120  
 QY 288 ORLWVEGORATLRCRCLGDPAPPTMHWGPPDDLVCNNSSRARAFPNGTLEIGVTGAGDAG 347  
 Db 121 ORLWVEGORATLRCRCLGDPAPPTMHWGPPDDLVCNNSSRARAFPNGTLEIGVTGAGDAG 180  
 QY 348 GYTCIATNPAGEATARVELRVIALPHGNSSAEGRPBPDIAASARTAAEGGTLESEP 407  
 Db 181 GYTCIATNPAGEATARVELRVIALPHGNSSAEGRPBPDIAASARTAAEGGTLESEP 240  
 QY 408 AVQVETVATSSGLVSNWGPGRPADPVMFQIQYNSSDETLIRIVASSHHEFLKHLVPG 467  
 Db 241 AVQVETVATSSGLVSNWGPGRPADPVMFQIQYNSSDETLIRIVASSHHEFLKHLVPG 300  
 QY 468 ADVDLCLALSPAAGSDTLATRLGCAHFTLPASPLCHALQAHVLCGLTVAVGVLY 527  
 Db 301 ADVDLCLALSPAAGSDTLATRLGCAHFTLPASPLCHALQAHVLCGLTVAVGVLY 360  
 QY 528 AALLVFTVALVGRGAGNRLPLKLSHVOSQTNGBSPTPKXHPSPSPPPPPQSCSLD 587  
 Db 361 AALLVFTVALVGRGAGNRLPLKLSHVOSQTNGBSPTPKXHPSPSPPPPPQSCSLD 420  
 QY 588 LGDAGCYGVARRLGAMARRSHVHGSLGAGRGVGSABERLEESVY 635  
 Db 421 LGDAGCYGVARRLGAMARRSHVHGSLGAGRGVGSABERLEESVY 468

## RESULT 8

US-09-732-436-24  
 ; Sequence 24, Application US/09732436  
 ; Publication No. US20030064919A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Prayaga, Subhidas A  
 ; APPLICANT: Shimkets, Richard K  
 ; TITLE OF INVENTION: Novel Polypeptides and Polynucleotides Encoding Same  
 ; FILE REFERENCE: 15966-615  
 ; CURRENT APPLICATION NUMBER: US/09/732,436  
 ; CURRENT FILING DATE: 2001-09-13  
 ; PRIOR APPLICATION NUMBER: 60/169,887  
 ; PRIOR FILING DATE: 1999-12-09  
 ; PRIOR APPLICATION NUMBER: 60/170,230  
 ; PRIOR FILING DATE: 1999-12-10  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 24  
 ; LENGTH: 832  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-732-436-24

Query Match 49.8%; Score 1662.5; DB 10; Length 832;  
 Best Local Similarity 46.4%; Pred. No. 1,3e-11;  
 Matches 364; Conservative 83; Mismatches 178; Indels 159; Gaps 13;

QY 9 LMASGA-----ACPLVCQCNLSSESLTCAHGLLFPVNVDRRTVEPLADNFIQAL 63  
 Db 51 LLAFGMAFAVVDACPKYCVCONLSBSLGTLCPSKGLFVPPDIDRTVEPLRNGNFIHT 110  
 QY 64 GPPFRFMTGLVDTLTLSRNATIRIGARAFGLDLSRLSHLDGNRLVELIGTSLGPNVLQ 123  
 Db 111 SRQDFANMTGLVDTLTLSRNATIRIGARAFGLDLSRLSHLDGNRLVELIGTSLGPNVLQ 170  
 QY 124 HLILSGNQLRIAGAFDDPLESLEDLDLSYNNLRQVFWAGIGAMPALHTINDHNTIDA 183  
 Db 171 HLIVNNNGGIGIADEAFEDFLTLEDLDLSYNNLHGHPWDSVVRMWNLHQSLDHNLDH 230  
 QY 184 LPPGAFQQLGSLDLTNSRLATLADPDLFSRGRDA--EASP--APVLSFSGNPLHNCN 239  
 Db 231 IABGTAFDLQKLARIADLTLSNRLQKLPPDPPIPARSOASALTATPAPPLSFGNPLHNCN 290

QY 240 CELLMLRLRLARPDDLETCA SPPLAGRYFMAVPEGEFSCEPPLIARTQRLWVLEGORAT 299  
 Db 291 CELLMLRLRLARPDDLETCA SPPLAGRYFMAVPEGEFSCEPPLIARTQRLWVLEGORAT 350  
 QY 300 LRCRAIGDPAFTMHWGPPDDLVCNNSSRARAFPNGTLEIGVTGAGDAGYTCIANTPAE 359  
 Db 351 LRCRAIGDPAFTMHWGPPDDLVCNNSSRARAFPNGTLEIGVTGAGDAGYTCIANTPAE 410  
 QY 360 ATARVELRVIALPHGNSSAEGRPBPDIAASARTAAEGGTLESEP-----AVO 410  
 Db 411 ATARVELRVIALPHGNSSAEGRPBPDIAASARTAAEGGTLESEP-----AVO 470  
 QY 411 VTEVATSSGLVSNWGPGRPADPVMFQIQYNSSDETLIRIVASSHHEFLKHLVPGADY 470  
 Db 471 VTEVATSSGLVSNWGPGRPADPVMFQIQYNSSDETLIRIVASSHHEFLKHLVPGADY 530  
 QY 471 DLCLALSPAAGSDTLATRLGCAHFTLPASPLCHALQAHVLCGLTVAVGVLYAL 530  
 Db 531 DLCLALSPAAGSDTLATRLGCAHFTLPASPLCHALQAHVLCGLTVAVGVLYAL 589  
 QY 531 LVFTVALVGRGAGNRLPLKLSHVOSQTNGBSPTPKXHPSPSPPPPPQSCSLD 577  
 Db 590 LVFTVALVGRGAGNRLPLKLSHVOSQTNGBSPTPKXHPSPSPPPPPQSCSLD 648  
 QY 578 -----PP----- 580  
 Db 649 NELLDTASLARASDSSSSSLGSGEAGLGRAPWRIPSPAPRPKPSLDRLMGAFASLDL 708  
 QY 581 -----ORQCSLD 587  
 Db 709 KSGRKEELDSRTPAAGAGTSARGHSREPLGPAPARASLDLPLEGKAKRSHSFD 768  
 QY 588 LGD-----AGCYGVARRLGAMARRSHVHGSLG-----AGCRGVGSABERLE 631  
 Db 769 MDPFMAAAGVPGGYSPPRKXSNITWTKSLSVNGMLPFEEESDLVGARGTGSSEPMW 828

## RESULT 9

US-10-004-378A-41  
 ; Sequence 41, Application US/10004378A  
 ; Publication No. US20030228301A1  
 ; GENERAL INFORMATION:

APPLICANT: Li, Li  
 APPLICANT: Furtak, Kazaryna  
 APPLICANT: Perna, Amanda  
 APPLICANT: Patutajan, Meera  
 APPLICANT: Shimkets, Richard A  
 APPLICANT: Guo, Xiaojia Saasha  
 APPLICANT: Casman, Stacie J  
 APPLICANT: Burgess, Catherine E  
 APPLICANT: Malyankar, Uriel M  
 APPLICANT: Tchernev, Velizar T  
 APPLICANT: Vernet, Corinne A  
 APPLICANT: Spytek, Kimberly A  
 APPLICANT: Agee, Michele  
 APPLICANT: Rastelli, Luca  
 APPLICANT: Shenoy, Suresh G  
 APPLICANT: Grose, William M  
 APPLICANT: Alsobrook II, John P  
 APPLICANT: Lepley, Denise M  
 APPLICANT: Gerlach, Valerie  
 APPLICANT: Edinger, Schomir  
 APPLICANT: MacDougall, John R  
 APPLICANT: Peyman, John A  
 APPLICANT: Gunther, Erik  
 APPLICANT: Stone, David J  
 APPLICANT: Ellerman, Karen  
 APPLICANT: Gangolli, Esia A  
 ; TITLE OF INVENTION: No. US20030228301A1e1 Human Proteins, Polynucleotides Encoding Tr

```

; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 21402-179
; CURRENT APPLICATION NUMBER: US/10/004,378A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/242,882
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,765
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/300,206
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/242,789
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,768
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,767
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/243,622
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/273,047
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/243,591
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,950
; PRIOR FILING DATE: 2000-10-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-004-378A-41

```

```

Query Match      49.8%; Score 1662.5; DB 14; Length 832;
Best Local Similarity 46.4%; Pred. No. 1,3e-111;
Matches 364; Conservative 83; Mismatches 176; Indels 159; Gaps 13;

```

```

QY 9 LIAAGAA-----ACPLPCVQONTLSSTLCAHRLGFVPVVDRTVELRLADNFIOAL 63
DB 51 LIAFGAFVAVVADACPKYCCQMLSESLGTLCPKGLFVPPIDRRTVELRGNFIHI 110
QY 64 GPPDRNNTGLVDLTLSNNAIRIGARAFGOLSESLHLDONRIVEIGTGLRPVNIQ 123
DB 111 SRQDRANTGLVDLTLSNNTLSHIQFSPFLDESLSRLHLSNRPISGEEDTLRGLVNIQ 170
QY 124 HLISGNQLRIAPGAFDDFLESLDLSYNNLRQVPAAGAMPALHTTLNLDNLDA 183
DB 171 HLIVNNQLGGIADAEFDELTELDLSTYNNLHGLFWDSVRRWVNIHQSLDENLDDH 230
QY 184 LPPGAFQOLGSLRDLTNSNRLATLAPPPLFSRGDA--EASP-APVLVSFGNPLHCN 239
DB 231 IAEGLFADLOKLARLDLTSNRLQKLPDPPIFARQASALATPAPPLSFGNPLHCN 290
QY 240 CELMLRLRLAPDDLETGASPPGLAGRYEMAVPEGESECEPPLIARHQLRWLEGGQAT 299
DB 291 CELMLRLERDDDETGSGPGLGRYFMVREBEFCEPPLIQTHTKLLVLEGQAT 350
QY 300 LRCRALGDPAPTMWVGDDRLVGNSSRRARAFPNGTLEIGYTGADGAGYCIATNPGE 359
DB 351 LKCAIGDPSPLIHVVAPDDRLVGNSSRTAYVNDTLITFTTSDSACFICIANANGE 410
QY 360 ATARVELRVALLPHGSSNAGGRGP--SDIASARTAAEGGTLESEP-----AVQ 410
DB 411 ATAWVEVSIQLPHLSNSTSRTPAPKSLDITGSSKTSRGGSAGGGEPPKSPPERAVL 470
QY 411 VTEWTATGVLVSWGGRPADPYMMQIOYNSSEDETLIRVVPASSHFLKHLVPGADY 470
DB 471 VSEVTTTSALVKKSVSAPRVKMTQLQYCSDEVLIRMI PANSKAFVNNNIVSGGY 530
QY 471 DLCLLAISPAAGPDLTATRLLGCAHFSTLPASPLCHALQAVLGTLTVAVGVLVVAL 530
DB 531 DLCLVLAAMDPTA-TTLTATNIVGCAQFETKADYDQCQSMHSQLIGTMILVIGGIYATL 589

```

```

QY 531 IVEFTVALLVYRGAGNGRPLKLU-----SHVQSGTNGG-----PSFTPKAHPRESP----- 577
DB 590 IVEFTVILMTRYK-VCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPGPKVYVR 648
QY 578 -----PRP----- 580
DB 649 NELDFTASLARASDSSSSSSSGSGEAAGLGRAPWRIPPSARPKSLDRIMGAFASLDL 708
QY 581 -----ORCSLD 587
DB 709 KQQRKEILLDSRTPAAGGAGTSARGHSDREBPLGPPAARASLLPLPLEGKAKRSHSFD 768
QY 588 LQD-----AGCYGARLGGAMARSHSVHGGILG-----AGCRGVGSGARLE 631
DB 769 MGDFPAAAGGVVPGGYSPPRKVSNITWKRSLSVNGMLPFRESDLVGARGTFSSEWVW 828
QY 632 ESVV 635
DB 829 ESTV 832

```

## RESULT 10

```

US-10-004-378A-42
; Sequence 42, Application US/10004378A
; Publication No. US20030228301A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Futak, Kazaryna
; APPLICANT: Perna, Amanda
; APPLICANT: Patutajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Casman, Stacie J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Vernet, Corinne A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Agee, Michele
; APPLICANT: Rastelli, Luca
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Grosse, William M
; APPLICANT: Alsebrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie
; APPLICANT: Edinger, Schlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Peyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangolli, Esra A

```

```

; TITLE OF INVENTION: No. US20030228301A1 Human Proteins, Polynucleotides Encoding T
; FILE REFERENCE: 21402-179
; CURRENT APPLICATION NUMBER: US/10/004,378A

```

```

; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/242,882
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,765
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/300,206
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/242,789
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,768
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,767
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/243,622
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/273,047
; PRIOR FILING DATE: 2001-03-02

```



```

; PRIOR APPLICATION NUMBER: 60/243,591
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,950
; PRIOR FILING DATE: 2000-10-27
; Remaining prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 42
; LENGTH: 789
; TYPE: PR
; ORGANISM: Cynomolgus monkey
US-10-004-378A-42

Query Match      49.7%; Score 1657.5; DB 14; Length 789;
Best Local Similarity 46.3%; Pred. No. 2.8e-111;
Matches 363; Conservative 82; Mismatches 180; Indels 159; Gaps 13;

QY 9 LIAAGAA-----ACPLPCVCONSESLSTLCARGLLFPVPVDRRTVELRLADNFIQAL 63
DB 8 LIAFGMAFAVVDACPKYCVCQONSESLGTLCPKGLFPFPDIDRTVELRLGNGFIHI 67
QY 64 GPPDFNNMTGLVDLTLSRNAITRIGARAFGDLSELSLHLDGNRLVELGTSLRGPNVLIQ 123
DB 68 SRQDFNNMTGLVDLTLSRNTISHIQFSLDELSLSLHLDNRSLGDLTGLVNIQ 127
QY 124 HLISGNQIGRIAPGAFDDFLESLELDLSYNNLRQVPWAGIGAMPALHTNIDNHLIDA 183
DB 128 HLIVNNQIGIADAEFEDELTLTLEDLDSYNNLHGLPMDSVARMVNLHQLSIDHNLIDH 187
QY 184 LPPGAFAOIGQSLRLDLTSNRALTATLAPDPLFSRGRDA--EASP--APVLVSFGNPLHCN 239
DB 188 IAEGTADLOKARLDLTSNRLOKLPDPPIFASQASALTATPAPPLSFGNPLHCN 247
QY 240 CELLMRLRLARPDDLETCSPPGLAGRYFWAVEGEFSCPEPLIARHTORLWVLEQORAT 299
DB 248 CELLMRLRLERDDLETCSPPGLKGRYFWAVEEVEFCPEPLITQHTHLVLBGOAAT 307
QY 300 LRCRALGDPAPTMWVGPDDRIVGNSSRAAFPNGLTEIGVTGAGAGYTCIATNPAGE 359
DB 308 LKCKALGDPSPPLIHWVAPDDRLVGNSSRTAVYNGTDFITTSQDSGATCIAANAGE 367
QY 360 ATARVELRLVALPHGNGSSAEGRPGP--SDIAASATAAEGTLESEP-----AVQ 410
DB 368 ATTVEVSTIQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGSGGGEPPKSPPERAVL 427
QY 411 VTEVATSGLVSWGPGRPADPVMWFOIQYNSSEDETLIYRIVPASSHFLKLHVPGADY 470
DB 428 VSEVTTTSALAKWSVSKSTPRVMYQIQYNCSDDEVLIYRMIPASNKAFFVNNLVSGTGY 487
QY 471 DLCLLALSPAAGSDLTATRLGCAHSTLPASPLCHALAHVUGTLTVAAGVLYAAL 530
DB 488 DLCLVLAAMWDDTA--TLTATATNIVGCAQFTKADYPOCCSMHSQILGTMILVIGIIVATL 546
QY 531 LVFTVALLVGRGAGNRLPLKL-----SHVOSQTNNG-----PSPTPKAHPRSP 577
DB 547 LVFTVILMVYK--VCNHEAPSKMAAAVSNVYSQTNNGAPPPSSAPAGAPPGPKVYVR 605
QY 578 -----PRP----- 580
DB 606 NEILDFTASLAPASDSSSSSLGSGAAAGLGRAPWRLPSPAPPKPSLDRLMGAFASLDL 665
QY 581 -----ORSCSLD 587
DB 666 KSKQKEELDLSRTPPAGGACTSARGHSDREPLIGPPAARASLLPLPLBCKAKKSSFD 725
QY 588 LGD-----AGCYGARRLGAMARRSHVHGGILG-----AGCRGVGGSABRL 631
DB 726 MGDFAAALAGGVVPGISPPRVSINMTKXSLSTVNGMLPPEESDVGAGGTPESSWVM 785
QY 632 ESVV 635
DB 786 ESTV 789

```

```

RESULT 11
US-10-291-172-261
; Sequence 261, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 261
; LENGTH: 789
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-291-172-261

```

```

Query Match      49.6%; Score 1655.5; DB 14; Length 789;
Best Local Similarity 46.3%; Pred. No. 3.9e-111;
Matches 363; Conservative 83; Mismatches 179; Indels 159; Gaps 13;

```

```

QY 9 LIAAGAA-----ACPLPCVCONSESLSTLCARGLLFPVPVDRRTVELRLADNFIQAL 63
DB 8 LIAFGMAFAVVDACPKYCVCQONSESLGTLCPKGLFPFPDIDRTVELRLGNGFIHI 67
QY 64 GPPDFNNMTGLVDLTLSRNAITRIGARAFGDLSELSLHLDGNRLVELGTSLRGPNVLIQ 123
DB 68 SRQDFNNMTGLVDLTLSRNTISHIQFSLDELSLSLHLDNRSLGDLTGLVNIQ 127
QY 124 HLISGNQIGRIAPGAFDDFLESLELDLSYNNLRQVPWAGIGAMPALHTNIDNHLIDA 183
DB 128 HLIVNNQIGIADAEFEDELTLTLEDLDSYNNLHGLPMDSVARMVNLHQLSIDHNLIDH 187
QY 184 LPPGAFAOIGQSLRLDLTSNRALTATLAPDPLFSRGRDA--EASP--APVLVSFGNPLHCN 239
DB 188 IAEGTADLOKARLDLTSNRLOKLPDPPIFASQASALTATPAPPLSFGNPLHCN 247
QY 240 CELLMRLRLARPDDLETCSPPGLAGRYFWAVEGEFSCPEPLIARHTORLWVLEQORAT 299
DB 248 CELLMRLRLERDDLETCSPPGLKGRYFWAVEEVEFCPEPLITQHTHLVLBGOAAT 307
QY 300 LRCRALGDPAPTMWVGPDDRIVGNSSRAAFPNGLTEIGVTGAGAGYTCIATNPAGE 359
DB 308 LKCKALGDPSPPLIHWVAPDDRLVGNSSRTAVYNGTDFITTSQDSGATCIAANAGE 367
QY 360 ATARVELRLVALPHGNGSSAEGRPGP--SDIAASATAAEGTLESEP-----AVQ 410
DB 368 ATAMVEVSTIQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGSGGGEPPKSPPERAVL 427
QY 411 VTEVATSGLVSWGPGRPADPVMWFOIQYNSSEDETLIYRIVPASSHFLKLHVPGADY 470
DB 428 VSEVTTTSALAKWSVSKSAPRVVMYQIQYNCSDDEVLIYRMIPASNKAFFVNNLVSGTGY 487
QY 471 DLCLLALSPAAGSDLTATRLGCAHSTLPASPLCHALAHVUGTLTVAAGVLYAAL 530
DB 488 DLCLVLAAMWDDTA--TLTATATNIVGCAQFTKADYPOCCSMHSQILGTMILVIGIIVATL 546
QY 531 LVFTVALLVGRGAGNRLPLKL-----SHVOSQTNNG-----PSPTPKAHPRSP 577
DB 547 LVFTVILMVYK--VCNHEAPSKMAAAVSNVYSQTNNGAPPPSSAPAGAPPGPKVYVR 605

```

QY 578 -----PRP----- 580  
Db 606 NELIDFTASLARSDDSSSSSLGSGEAGLGRAPWRIIPSPAPRPKPSLDRLMGAFASLDL 665  
QY 581 -----ORCSLD 587  
Db 666 KSQRKEELDSRTPAAGAGTSARGHHSREPLLGPPAPARASLLPLPLEGAKRSHSFD 725  
QY 588 LGD-----ACGYARRLGAGWARRSHSVHGLG-----ACGRVGGSARLE 631  
Db 726 MGFPAALAAAGVAVPGVSPPRKXSNWTKRSLSVNGMLLPFEESDLVGARGTGSSSEWVR 785  
QY 632 ESTV 635  
Db 786 ESTV 789

RESULT 12  
US-10-221-278-261  
; Sequence 261, Application US/10221278  
; Publication No. US20040034208A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-045  
; CURRENT APPLICATION NUMBER: US/10/221,278  
; PRIOR FILING DATE: 2002-09-06  
; PRIOR APPLICATION NUMBER: 09/693,267  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 09/665,363  
; PRIOR FILING DATE: 2000-09-19  
; PRIOR APPLICATION NUMBER: 09/616,847  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 09/596,193  
; PRIOR FILING DATE: 2000-06-17  
; PRIOR APPLICATION NUMBER: 09/574,454  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: 09/519,705  
; PRIOR FILING DATE: 2000-03-07  
; NUMBER OF SEQ ID NOS: 752  
; SEQ ID NO 261  
; LENGTH: 789  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-221-278-261

Query Match 49.6%; Score 1655.5; DB 15; Length 789;  
Best Local Similarity 46.3%; Pred No. 3.9e-111;  
Matches 363; Conservative 83; Mismatches 179; Indels 159; Gaps 13;

QY 9 LIASGAA-----ACPLPCVQNTLSSTLCAHRLGLFVPPVNDRTVELRLADFIQAL 63  
Db 8 LIAFGMAFAVVDACPKYCCVQNTLSSTLGLCPKSLVPPDIDRITVELNLGNFIHI 67  
QY 64 GPPFRMNTGIVDLTSLNATRTICARFPGDLSEIRSHLDGNRLVELGSLRGPVNLQ 123  
Db 68 SRDPANNTGLVDLTSLNITSHIQPSFLDESIRSHLDGNRLVELGSLRGPVNLQ 127  
QY 124 HLISGNQIGRIAPGAFDFLESLEDDLTSTNNLRQVWAGIGAMPALHTLNLHNLIDA 183  
Db 128 HLINNNQIGRIADAEFDELITLEDLSTNNLHGLWDSVRRVNLHQLSLDNLHLDH 187  
QY 184 LPPGAFQGLSLRLDLSNRLATLAPDLPSRGRDA--EASP--APLVLSFGSNPLHCN 239  
Db 188 IAEGTFAQLKTLARLDLTSLNRLQTLPPPIFARQASALTLTPAPPLSEFGSNPLHCN 247  
QY 240 CELLMRLIARPDULETCASPPGILAGRYFAVPPGSGCEPPLIARHTQRLVLEGGAT 299  
Db 248 CELLMRLIARPDULETCASPPGILAGRYFAVPPGSGCEPPLIARHTQRLVLEGGAT 307  
QY 300 LRCRALDPAFTMHWGDDRLVGNSSRARAFPNGTLEIGVTAGDGAGYTCIATNPAGE 359

Db 308 LKCAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTPDIFITTSQDSGAFCTIAANAGE 367  
QY 360 ATARVRLRLIALPHGNSNAEGGRPGP--SDIAASARITAEGETLESEP-----AVQ 410  
Db 368 ATAMVEVSIVOLPHELNSTSRTPAPKRSRLSDITGSSKTSRGSGSGGGEPPKPPERAVL 427  
QY 411 VTEVATSGIVSGPGRPADPVMWFOIYVNSSDELTLYRIYVASSHPLIKLVLPGADY 470  
Db 428 VSEVITTSALIVKMSVSKAPRVKMTQLYNCSDEVLILYKMIIPASKAFVNNLVSGTGY 487  
QY 471 DLCLALSPAAGPSDLTATRLIGCAHFTLTPASPLCHALQAVHGLTTLVAVGVVYAL 530  
Db 488 DLCLAMWDDTA--TLTATNIVGCAQFTFKADYPPQGSVHSQTLGSTMILVIGIIVATL 546  
QY 531 LVFTVALIVRGAGNGRLPLKL-----SHVQQTNG-----PSPPKAPPPRSP 577  
Db 547 LVFTVILWVYK--VCNHEAPSKMAAIVSNVQTMGACPPPPSAPAPAPQGPFRVVR 605  
QY 578 -----PRP----- 580  
Db 606 NELIDFTASLARSDDSSSSSLGSGEAGLGRAPWRIIPSPAPRPKPSLDRLMGAFASLDL 665  
QY 581 -----ORCSLD 587  
Db 666 KSQRKEELDSRTPAAGAGTSARGHHSREPLLGPPAPARASLLPLPLEGAKRSHSFD 725  
QY 588 LGD-----ACGYARRLGAGWARRSHSVHGLG-----ACGRVGGSARLE 631  
Db 726 MGFPAALAAAGVAVPGVSPPRKXSNWTKRSLSVNGMLLPFEESDLVGARGTGSSSEWVR 785  
QY 632 ESTV 635  
Db 786 ESTV 789

RESULT 13  
US-10-004-378A-43  
; Sequence 43, Application US/10004378A  
; Publication No. US20030228301A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Li  
; APPLICANT: Furtak, Kazazyna  
; APPLICANT: Petura, Amanda  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Guo, Xiaojia Sasna  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Malyankar, Urfiel M  
; APPLICANT: Tchervet, Veizar T  
; APPLICANT: Vernet, Corinne A  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Agee, Michele  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Grose, William M  
; APPLICANT: Alsbrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Edinger, Scholmit  
; APPLICANT: MacDougall, John R  
; APPLICANT: Peyman, John A  
; APPLICANT: Gunther, Erik  
; APPLICANT: Stone, David J  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Gangoli, Esha A  
; TITLE OF INVENTION: No. US20030228301A1el Human Proteins, Polynucleotides Encoding T  
; FILE REFERENCE: 21402-179  
; CURRENT APPLICATION NUMBER: US/10/004,378A  
; PRIOR FILING DATE: 2001-10-24  
; PRIOR APPLICATION NUMBER: 60/242,882

PRIOR APPLICATION NUMBER: 60/242,765  
 PRIOR FILING DATE: 2000-10-24  
 PRIOR APPLICATION NUMBER: 60/300,206  
 PRIOR FILING DATE: 2001-06-22  
 PRIOR APPLICATION NUMBER: 60/242,789  
 PRIOR FILING DATE: 2000-10-24  
 PRIOR APPLICATION NUMBER: 60/242,768  
 PRIOR FILING DATE: 2000-10-24  
 PRIOR APPLICATION NUMBER: 60/242,767  
 PRIOR FILING DATE: 2000-10-24  
 PRIOR APPLICATION NUMBER: 60/243,622  
 PRIOR FILING DATE: 2000-10-26  
 PRIOR APPLICATION NUMBER: 60/273,047  
 PRIOR FILING DATE: 2001-03-02  
 PRIOR APPLICATION NUMBER: 60/243,591  
 PRIOR FILING DATE: 2000-10-26  
 PRIOR APPLICATION NUMBER: 60/243,950  
 PRIOR FILING DATE: 2000-10-27  
 Remaining Prior Application data removed - See file Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 191  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 43  
 LENGTH: 788  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-10-004-378A-43

Query Match 48.9%; Score 1630; DB 14; Length 788;  
 Best Local Similarity 46.0%; Pred. No. 2,7e-109;  
 Matches 360; Conservative 84; Mismatches 181; Indels 158; Gaps 13;

QY 9 LLAGAA-----ACPLPCVQCNLSSTLTCARGLFVPPNDRRTVELRLADNFIOAL 63  
 DB 8 LIAFGAFAVAVDACPXYCVCQNLSESLGTLCPKRLFPVPPDRLRVELRGLNFIHH 67  
 QY 64 GPPDPNNMGLVLTLSRNAITRIGARAFGLDSLRLHDGRVLELGTSLRGVNLQ 123  
 DB 68 GRODFANMGLVLTLSRNITSHIQPSFLDLSLSLHDSKRLSLGSDTLRGVNLQ 127  
 QY 124 HLISGNOLGRIAGAFDDFLESLEDLDSYNNLRQVPMAGIGAMPALHTLNDHNLDA 183  
 DB 128 HLIVNNQGLIGDADAEEDFLTLEDLDSYNNLHGLPMDSVRRMVLHQLSDHNLDDH 187  
 QY 184 LPPGAFAQLGSLRLDLSNRATLAPDLFSRGRDA--EASP--APLVSESGNPLHCN 239  
 DB 188 IAGTFADLOKLRLDLTSNRLOKLPDPIFASQASLLTATPAPPLSFGGNPLHCN 247  
 QY 240 CELLMLRRLARPDDLETCASPPGLAGRYFWAVGEGFSCPEPLIARTORIWLVEGORAT 299  
 DB 248 CELLMLRRLERDDLETCGSPGSLKGRYFWHIREEFVCEPPLITQHTHKLVLVEGOAT 307  
 QY 300 LRCRALDPAPTMHWGPDRLVGNSSRAAPFNGTLEIGVTAGDAGYTCIATNPAGE 359  
 DB 308 LKCAIGDPSPLIHWVAPDRLVGNSSRTAVYDNGTLDIITTSQDSGPTCIAANNAAGE 367  
 QY 360 ATARVELRVALPFGNSSAEGGRPG--SDIAASARTABEGSTLESEP-----AVQ 410  
 DB 368 ATATVESIVQLPPLSNSTSRMAPPKSRLSDITGSSKTSRGGGSGAGEPEKSTPEPAVL 427  
 QY 411 VTEVTATSGLVNMGPRPADPVMWFOIQYNSSEDETLIYRIVPASHHFLKHLVPGADY 470  
 DB 428 VSDVTTTSAIVKSVKSAFVVMYQOYNCSDDVILYIMIPASNAFVYNNLVSGTGY 487  
 QY 471 DLCLALSPAAGPSDLTATRLGCAHFSTLPASPLCHALQAHVIGTLTVAAGVLYAAL 530  
 DB 488 DLCLALSPAAGPSDLTATRLGCAHFSTLPASPLCHALQAHVIGTLTVAAGVLYAAL 546  
 QY 531 LVFVALLVGRGNGRLPLKL-----SHVQSQTNGGSP-----TPKXHP--- 573  
 DB 547 LVFVILLVGRYK--VCNHDTPGKMAAATVSNVYSQTSQSPPLGIPVQQLPQAPKVV 605  
 QY 574 -----RSPPPR----- 580

DB 606 RNEIMDFSTSLARACDSSSSSLSGSEAAHGRGWRLPPAPRKPSLDRLMGAFASID 665  
 QY 581 -----ORCSL 586  
 DB 666 LKSORKEBLDSRTPAGAGCTSSRGHSHDREPLIGPATRASLLPLLEBKAKRSHSF 725  
 QY 587 DLGD-----AGCYGARRLGAMARRSHSHVHGLIG-----AGCRGVGSAERLEE 632  
 DB 726 DWGDFAAAAAIVPGYSPPRRVSNITWTRKRSYNNMLLPFERSDLVGARGTGSESMVWE 785  
 QY 633 SVV 635  
 DB 786 STV 788

RESULT 14  
 US-09-815-626-2  
 Sequence 2, Application US/09815626  
 Patent No. US2002076752A1  
 GENERAL INFORMATION:  
 APPLICANT: Gluckmann, Maria Alexandra  
 TITLE OF INVENTION: 33395, A NOVEL HUMAN LEUCINE-RICH REPEAT  
 TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF  
 FILE REFERENCE: 10448-031001  
 CURRENT APPLICATION NUMBER: US/09/815,626  
 CURRENT FILING DATE: 2001-03-23  
 PRIOR APPLICATION NUMBER: US 60/191,863  
 PRIOR FILING DATE: 2000-03-24  
 NUMBER OF SEQ ID NOS: 17  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 2  
 LENGTH: 628  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-815-626-2

Query Match 46.7%; Score 1557; DB 9; Length 628;  
 Best Local Similarity 54.0%; Pred. No. 4e-104;  
 Matches 329; Conservative 71; Mismatches 175; Indels 34; Gaps 8;

QY 4 PLILILLAGAAA-----CPPLPCVQCNLSSTLTCARGLFVPPNDRRTVEL 53  
 DB 5 PLILILPLAPASSPPQSATPSPBCPRRCQOTQSLPSVLCPPAGALLFPSPSLDRPAEL 64  
 QY 54 RLADNFIOALGPPDPNNMGLVLTLSRNAITRIGARAFGLDSLRLHDGRVLELGT 113  
 DB 65 RLADNFITASRRDLANMGLHLSRNITRIVAAAFADRLALALHDGRVLSIGE 124  
 QY 114 GSLRGVNLQHLISGNOLGRIAGAFDDFLESLEDLDSYNNLRQVPMAGIGAMPALHT 173  
 DB 125 GQLRGVNLRLHLSNNQLAALAAAGALDDCAETLEDLDSYNNLEQLPWEALGRIGVNT 184  
 QY 174 LINDHNLIDALPPGAFAQLGSLRLDLSNRATLAPDLFSR-----GRDASAPAVLV 229  
 DB 185 LGIDHNLASVPAAGARSRLKRLALDWTNRLLTTPIDPLFSRLPLARRGSPASALVL 244  
 QY 230 SFGSNPLHNCCELLMLRRLARPDDLETCASPPGLAGRYFWAVGEGFSCPEPLIARTOR 289  
 DB 245 AFGGNPLHNCCELLMLRRLAREDDLETCASPPGLAGRYFWAVGEEFVCEPPLVTHRSP 304  
 QY 290 LMTLEGORATLRCRALDPAPTMHWGPDRLVGNSSRAAPFNGTLEIGVTAGDAGY 349  
 DB 305 LAVPAGRPALRCRAVDDPEPRVWVSPQGRLLGNSSRAAPFNGTLELVTBEGDGIGIF 364  
 QY 350 TCIATNPAGATARVELRVALHP--CGNSSAEGGRPG-----PSDIAASARTAAEG 399  
 DB 365 TCIANAAGAGATAVELTVBPPEPQLANSTCDPPDGPDPALTPPSAASAKAVDTG 424  
 QY 400 EGTLESPAVQVTEVATSGLVNMGPRPADPVMWFOIQYNSSEDETLIYRIVPASHHF 459  
 DB 425 PPT---DRGYQVTEHGATLALVQWPDORPIPIGRMYIOIQYNSADDILVYRMIPAESRSF 481  
 QY 460 LKHLVPGADYDLCIALSPAAGPSDLTATRLGCAHFSTLPASPLCHALQAHVIGTLT 519

Db 482 LUTDLASGRTYDLCVLAIVEGSA-TGLTATRPVGARFSTEPALRPGAHAFLLGGTMI 540  
QY 520 VAVGVIVAAALVFTVALIVR-----GRGANGRLPLKXSHVOSQTNGSPPTPKAHPR 574  
Db 541 IALGVIVASVIVFIVFLMRKYKHGGQPPGAKIIPAPVSSVCSQTNMGALGFTPTAPP- 599  
QY 575 SPPEPRQRS 583  
Db 600 AEPAPALRA 608

## RESULT 15

US-10-145-586-61  
Sequence 61, Application US/10145586  
Publication No. US20030138890A1  
GENERAL INFORMATION:  
APPLICANT: Alexandra Gluckemann, Maria  
APPLICANT: Silos-Santiago, Immaculada  
APPLICANT: M. Galvin, Katherine  
APPLICANT: Weich, Nadine  
APPLICANT: Curtis, Rory A.J.  
APPLICANT: Bandaru, Rajasekhar  
APPLICANT: Kapeller-Liebermann, Rosana  
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTOR FAMILY MEMBERS,  
TITLE OF INVENTION: HUMAN THIOREDOXIN FAMILY MEMBERS, HUMAN LEUCINE-RICH  
FILE OF INVENTION: REPEAT FAMILY MEMBERS, AND HUMAN RINGFINGER FAMILY MEMBER  
FILE REFERENCE: 10448-188001  
CURRENT APPLICATION NUMBER: US/10/145,586  
CURRENT FILING DATE: 2002-05-14  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 95  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 61  
LENGTH: 628  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-145-586-61

Query Match 46.7%; Score 1557; DB 14; Length 628;  
Best Local Similarity 54.0%; Pred. No. 4e-104;  
Matches 329; Conservative 71; Mismatches 175; Indels 34; Gaps 8;

QY 4 PULLILLASGAAA-----CPUCVQNLSESLSLTCARHGLFPVNVDRRTVEL 53  
Db 5 PULLCLPLAPASSPPOSATPSPPCRRCQCTQSLPLSVLCPGAGLFPVPSLDRRAEL 64  
QY 54 RLADNFIOALGPPDRNMVGLVDLTLSRNATIRIGARAFGOLESRLHLDGNRLVELGT 113  
Db 65 RLADNFIVASVRRDLANMTGLHLULSLSRNTRIVAAAFADRLALRLHLDGNRLTSLGE 124  
QY 114 GSLRGPNVLOHLIISGNQLGRIAPGAFDDFLESLEDLDSYNNLRQVPMAGIGAMPAH 173  
Db 125 GOLRLVNLRLHLLISNNQLAALAGALDDCAETLEDLDSYNNLEQLPMEALGRIGNVT 184  
QY 174 LNLBNLIDALPPGAFQOLGQSLRDLTSNRLATTAPDLFSR---GRDAASAPAPVL 229  
Db 185 LGLBNLLASVPAQFSLRHLARLDMTSNRLTTPDPLFSRLPLARPRGSPASALVL 244  
QY 230 SFGNPLHNCCELMLRLARPDDLETGASPRGLAGRYFMAVPEGEFCEPPLIARHTQR 289  
Db 245 AFGNPLHNCCELMLRLARDDLEACASPPALGGRYFMAVGEFEFVCEPPEVYTHRSPP 304  
QY 290 LMTLEGQATLRCALGDPAPTMHWVGPDDRLVGNSSRARAFPNGLTLEIGVTGAGDAGY 349  
Db 305 LAVPAGRPALRCRAVGDPPEPRVHRVSPQGRLLGNSSRARAFPNGLTLELVTEPDGIF 364  
QY 350 TCITATNPAGETAKAYELRYALPH---GNSASAEGRPG-----PSDIAASARTAAEG 399  
Db 365 TCIAANAAGEATYAAVELTVGPPPPQLANSTSCDPPRPDDDALTPPAAASAKVADTG 424  
QY 400 EGTLESEPAVQVTEVATSGLVSWGPRPADPVMMFOIQVNSSEDETLIYRIVPASSHIF 459

Db 425 PPT---DRGVQVTEHGATAALVQMPDQRPPIGIRMTQIQVNSSADDILYVRMIPAESRSF 481  
QY 460 LKHLVPGADYDLCILALSPAGPSDLTATRLILGCHFFSTLPASPILCHALQAHVLGSTLT 519  
Db 482 LUTDLASGRTYDLCVLAIVEGSA-TGLTATRPVGARFSTEPALRPGAHAFLLGGTMI 540  
QY 520 VAVGVIVAAALVFTVALIVR-----GRGANGRLPLKXSHVOSQTNGSPPTPKAHPR 574  
Db 541 IALGVIVASVIVFIVFLMRKYKHGGQPPGAKIIPAPVSSVCSQTNMGALGFTPTAPP- 599  
QY 575 SPPEPRQRS 583  
Db 600 AEPAPALRA 608

Search completed: November 17, 2004, 22:52:27  
Job time : 145 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 22:34:49 ; Search time 39 Seconds  
(without alignments)  
1079.793 Million cell updates/sec

Title: US-10-071-879-10

Perfect score: 3336

Sequence: 1 MAPPLLLLLIAGSAAACPLP.....LGAGCRGVGSARLEESVY 635

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 segs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: /cgnt2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
2: /cgnt2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
3: /cgnt2\_6/prodata/1/iaa/6A\_COMB.pep:\*  
4: /cgnt2\_6/prodata/1/iaa/6B\_COMB.pep:\*  
5: /cgnt2\_6/prodata/1/iaa/pCtUS\_COMB.pep:\*  
6: /cgnt2\_6/prodata/1/iaa/backfill1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	425.5	12.8	713	4	US-09-907-794A-245
2	425.5	12.8	713	4	US-09-905-125A-245
3	425.5	12.8	713	4	US-09-902-775A-245
4	425.5	12.8	713	4	US-09-906-700-245
5	425.5	12.8	713	4	US-09-903-603A-245
6	415.5	12.4	670	3	US-09-063-950-2
7	407.5	12.2	640	4	US-09-907-794A-292
8	407.5	12.2	640	4	US-09-905-125A-292
9	407.5	12.2	640	4	US-09-902-775A-292
10	407.5	12.2	640	4	US-09-906-700-292
11	407.5	12.2	640	4	US-10-140-002-368
12	407.5	12.2	640	4	US-09-903-603A-292
13	386.5	11.6	590	4	US-09-520-781-12
14	385.5	11.6	745	4	US-10-140-002-68
15	383	11.5	620	4	US-09-907-794A-73
16	383	11.5	620	4	US-09-905-125A-73
17	383	11.5	620	4	US-09-902-775A-73
18	383	11.5	620	4	US-09-906-700-73
19	383	11.5	620	4	US-10-140-002-324
20	383	11.5	620	4	US-09-903-603A-73
21	382.5	11.5	653	4	US-10-140-002-438
22	378.5	11.3	653	4	US-09-520-781-10
23	374.5	11.2	448	4	US-09-520-781-32
24	359.5	10.8	598	4	US-09-866-028-69
25	359.5	10.7	598	4	US-09-944-457-69
26	356.5	10.7	716	4	US-10-140-002-512
27	351	10.5	428	4	US-10-140-002-292

28	332.5	10.0	708	3	US-09-131-648-2	Sequence 2, Appl
29	332.5	10.0	708	4	US-09-907-794A-69	Sequence 69, Appl
30	332.5	10.0	708	4	US-09-905-125A-69	Sequence 69, Appl
31	332.5	10.0	708	4	US-09-902-775A-69	Sequence 69, Appl
32	332.5	10.0	708	4	US-09-906-700-69	Sequence 69, Appl
33	332.5	10.0	708	4	US-10-140-002-298	Sequence 298, Appl
34	332.5	10.0	708	4	US-09-903-603A-69	Sequence 69, Appl
35	312.5	9.4	1101	3	US-08-986-485-2	Sequence 2, Appl
36	308	9.2	1091	3	US-08-986-485-5	Sequence 5, Appl
37	307	9.2	1059	4	US-09-907-794A-290	Sequence 290, Appl
38	307	9.2	1059	4	US-09-905-125A-290	Sequence 290, Appl
39	307	9.2	1059	4	US-09-902-775A-290	Sequence 290, Appl
40	307	9.2	1059	4	US-09-906-700-290	Sequence 290, Appl
41	307	9.2	1059	4	US-09-903-603A-290	Sequence 290, Appl
42	307	9.2	1119	4	US-09-907-794A-294	Sequence 294, Appl
43	307	9.2	1119	4	US-09-905-125A-294	Sequence 294, Appl
44	307	9.2	1119	4	US-09-902-775A-294	Sequence 294, Appl
45	307	9.2	1119	4	US-09-906-700-294	Sequence 294, Appl

## ALIGNMENTS

RESULT 1  
US-09-907-794A-245  
Sequence 245, Application US/09907794A  
Patent No. 6635468  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gertlisen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Grimaldi, Paul J.  
APPLICANT: Guiney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kjaevan, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/907,794A  
PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547

```

; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 245
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-907-794A-245

Query Match      12.8%; Score 425.5; DB 4; Length 713;
Best Local Similarity 25.2%; Pred. No. 1.4e-26;
Matches 182; Conservative 74; Mismatches 258; Indels 207; Gaps 26;

QY 1 MAPPILLILLASGAA-----CPLPCVCO-----NUSESTLCAHRLFLV 42
DB 4 LVAPLLIYAWAGATATVVPVHVPCPPQACQIRPMTYPRSSRRELTVDCCNDLFLITAV 63
QY 43 PNVVDRRTVELRLADNFICQAGPPFRMTGLVDLTLSRNAITRIGARFGLDLSRLH 102
DB 64 PPAIPATQTLILQSNISIVAVDQELGYLANLTFLDLSQNSFSDARCCDFFALPQLLSTH 123
QY 103 LDGRRVLELGTSLRGVNIQHLISNQGLIAPGAFD----- 141
DB 124 LERNQLTRLEDHSPAGLSIQELVLTNNQYRIAPRAFSGLSNLRHLNSNLRADISR 183
QY 142 -----DFLESLDLD 151
DB 184 WFEMLEPULIMTGNKVVDALIDMNPRLANSLVLAQMNLRIISDYALEGLQSLSL 243
QY 152 LSYNNLRQVWAGIGAMPALHTTLNLDHNLIDALPPGAF----- 190
DB 244 FYDQGLARVRRALAEQVPGKFLDLNKNPIQVGPFGFANMLHKLGLNNMELVSLDK 303
QY 191 ----QCGQLSRDLTSLN-RLATLAPDPLFS-----RGRDAKSPAPLV 228
DB 304 FALVNLPELTLDITNNPRLSFIPRAFHLPQMETIMNNALSALHQVLESLEPNIQE 363
QY 229 LSPSGNPLHCHNCCELLMLR-----RLARPDLFCASPPGLAGRYFAVPEGES--CEP 280
DB 364 VGLHGNPRCCVIRMANATGTRVRFTEPQS-TICAEPRDQIRPVAREVPREMTDHCIP 422
QY 281 PLIAR-HTORLMLVEGORATLRCRALGDPAPTMWVGDD-RLVGNSS--RARAPENGTL 336
DB 423 LISRSPSPSLQVASGESWVLHCRALAPBEIYVWTGAGRILTPAHAGRRYRVPEGTL 482
QY 337 EIGVTGAGDAGGYTCIATNPAGEATARVELAVLALPHOGNSAAGGRRGPDIAASRTA 396
DB 483 ELRVTAEAGLYTCVQNLVAGADTKTVSVV-----GRALLQPGRD----- 524
QY 397 AEGEGTLESEBPAVQVTEVATSGLVSW--GGRPADPVMMFOIQVNSEDETLTYRIYPA 454
DB 525 -EGGS-LR-----LVQETHPHYHILISWTPRYNTVSTNLTWSSASLSRGQATLALRL-ER 577
QY 455 SSHHFLKHLVPGADYDLCILALSPAAGPSDLTATRLIGCAHFSTLPASPLCHALQAHVL 514
DB 578 GTHSYNITRLILQATEVYMACIQV-----APADAHTQACVWARTKEATS-CHR----- 623
```

```

QY 515 GGLTYAVG---GVVMAILVFTVALLVGRGA--GNGRLPLKTSVQSOTNGSPSPTK 569
DB 624 -----ALDPRPLI--AIIALAVILLAGLAHLSTGQ-PRK-----GVGGRPLDP 667
QY 570 A 570
DB 668 A 668

RESULT 2
US-09-905-125A-245
; Sequence 245, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
```

```

; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 245
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-905-125A-245

Query Match      12.8%; Score 425.5; DB 4; Length 713;
Best Local Similarity 25.2%; Pred. No. 1.4e-26;
Matches 182; Conservative 74; Mismatches 258; Indels 207; Gaps 26;

QY 1 MAPPLLLLLLASSAAA-----CPPLVCQ-----NLSSSLTLCARGLLTV 42
DB 4 LVAPLLAMVAGATATVPVVPVWHPVPCPCQACQIRPMYTPRSSYREATTVDNDLPLTAV 63
QY 43 PNVDRRTVELRADNFIQALGPDPFRNMGTGLVDLTLSRAITRIGARAFGDELSRSLA 102
DB 64 PRLPAGTQTLILQSSIVRVDSSELYLANLTLELDSQNSFSDARDCHHALPQLLSLH 123
QY 103 LDGNRIVELGTGSLRGPVNLQHLILSNGOLGRIAPGAFD----- 141
DB 124 LERNQUTRLDEHSHFAGIASIQELYNHNLQYRIAPRAFGLSNLRLHNSNLRAIDSR 183
QY 142 -----DPLESLDD 151
DB 184 WERMENLILMIGKNKVDALLDMNFRPLANISLVLAGMNLREISDYALEGLQSLSSIS 243
QY 152 LSYNNLRQVPMWAGAMPALHTLNLDPNLIDALPQAPFA----- 190
DB 244 FYDNOGLARVPRRLLEGVPGKFLDKNPLQKRVGPDPFAMMLHKLGLANNEELVSTDK 303
QY 191 ----QIGQLSRDLTJN-RIATLAPDPLFS-----RGRDAEASAPPLV 228
DB 304 FALVNLPELTKLDTNNPRLSFTHPRAFHHLPMQETMLNNAALSHQGVESLNLQE 363
QY 229 LSFSGPRLNCNCLMLLR-----RLARPDULETCSFPGLAGRYWAVPEGEFS--CEP 280
DB 364 VGLHNPRLCDQVIRMANATGRVRFLEPOS--TLCAEPDLORLPVREVPFRMTDHCJP 422
QY 281 PLIAR-HTQRLWLVEGQATLRCRAGDPAPTWGWPDD-RLVGNSS--PARAPNGTL 336
DB 423 LISPRFPPSLQVASESMVLRHRALEBEPETIYWTTPAGLRLTPAHAGRRIRVYPEGTL 482
QY 337 ELGVTGAGDAGYTCIATNPAGEATARVELRYLALPHGNSAABGGRPGPSDIASARTA 396
DB 483 ELRRVTAEBEAGLYTCAQNLVGADTKTVSVVV-----GRALLQPERD----- 524
QY 397 AGEGETLESEPAVOUEVATSGLVSW--GPGRPADPVWMPFOIYNSSDEDTLYRIYVA 454
DB 525 -BQOG-LE---LRVQETHPHYHLLSWTPPMTVSNTLWSSASSLRQGAALALARTL-BR 577
QY 455 SSHHFLKLHPGADVDLCLIALSPAAGSDLTATRLLCAPHSSTLPASPLCAALQANVL 514
DB 578 GHHSVYITRLQLATEWACLQV-----AFADAHQLQALCWAARTKENS-CHR----- 623
QY 515 GGTLYAVAG--GVVLAALLFTVALLVGRGA--GNGRLPLKLSHVQQTNGGSPPTPK 569
DB 624 -----ALGDRPGLI--ALIALAVILLIAGLAALHGTGQ-PRK-----GVGGRPLLP 667
QY 570 A 570
DB 668 A 668
```

```

; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Bostein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurley, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 245
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-902-775A-245
```



Query Match 12.8%; Score 425.5; DB 4; Length 713;  
 Best Local Similarity 25.2%; Pred. No. 1.4e-26;  
 Matches 182; Conservative 74; Mismatches 258; Indels 207; Gaps 26;

QY 1 MAPPLILLILASGAAA-----CPLPCVCO-----NLSESLTLCARGLLFV 42  
 Db 4 LVAPLLIAMIWAGATATVPVPMVHPVPCPCQACQIRPMYTPRSSYRBAITVDCNDLFLTAV 63  
 QY 43 PPNVDRRTVELRLADNFIQALGPPDFRNMGTGLVDLTLSRAITRIGARAGDLESLSLH 102  
 Db 64 PPALPACTQTLLLOSNSIVAVDQSELGIANLTLELDLSQNSFSFARDCCDFHALPQLLSLH 123  
 QY 103 LDGNRLVELGTSLRGVNLQHLILSGNQGRIAPGAFD----- 141  
 Db 124 LEENQLTRELDHSPAGLASLOELYLNQLYRIAPRAFSGLSNILRLHNSNLRADSR 183  
 QY 142 -----DFLESLEDD 151  
 Db 184 WFEMLPVLITLMIGNKVDAILDKNFPPPLANLSVLAKNNLREISDYALEGLQSESLIS 243  
 QY 152 LSTVNNLRQVWAGIGAMPALHTLNDNLIDALPRGAF----- 190  
 Db 244 FYDNLARVPRRALEQVPGKFLDLNPNLQRVGPGDFAMMLHKLGLNNMELVSIDK 303  
 QY 191 -----QLGOLSRDLDTSN-RIATLAPDPLFS-----RGDAEASPAFLV 228  
 Db 304 FALVNLPELTIKDLITNNPRLSFIHPRAPHLPMQETMLNNNALSHQGVESLPLQE 363  
 QY 229 LSPSGNPLHCNCELLMIR-----RLARPDDLETQASPGIAGRYFAVPEGES--CEP 280  
 Db 364 VGLHGNIRDCDCVIRMANATGTRVRFTEPOS-TLCAEPDLQRLPVRVDFPREMTDHLCP 422  
 QY 281 PLIAR-HTQGLWVLEGORATLRCALGDPAFTMMWGPDD-RLVGNSS--RAAFEPGTL 336  
 Db 423 LISRSPSPSLQVASSGSMVHCRLAEPPEIYVWIPAGRLTPAHAGRYRYPGTL 482  
 QY 337 EIGVTGAGDAGGYCIAITNPAGEATARELEVLALPHGNSAAGRPGPSDINASARTA 396  
 Db 483 ELKRVTAEBAGLYTCVQONLVGADTKTVSVV-----GRALLQPGD----- 524  
 QY 397 AEGGTLESEPAVQVTEVATSGLVSW--GGRPADPVMWFQIOYNSSEDETLIYRIVA 454  
 Db 525 -EGGG-LE---LKVQETHPYHILLSWTPNTVSTMLTSSASLSRGQATALARL-PR 577  
 QY 455 SSHHFLKLHVPADYDLCIALSPAGPSDLTTRLLGCAHFTLPAAPLCHALQAHV 514  
 Db 578 GTHSYNTRILLQATEYACLOV-----APADAHQOLACVMARKETKS--CHR----- 623  
 QY 515 GGLTVAVG--GVIVALLVFTVALIVRGGA--GNGRLPLKLSHYOSQTNCGSPSPTPK 569  
 Db 624 -----ALGDRPGLI--ALLALVLLLAAGLAHLGTGQ-PRK-----GVGGRRLPLP 667  
 QY 570 A 570  
 Db 668 A 668

RESULT 4  
 US-09-906-700-245

; Sequence 245; Application US/09906700  
 ; Patent No. 6723535  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Genentech, Inc.  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Filvarcoff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, Christopher J.  
 ; APPLICANT: Guiney, Austin L.  
 ; APPLICANT: Hillan, Kenneth, J.  
 ; APPLICANT: Kijavlin, Ivar J.  
 ; APPLICANT: Mather, Jennie P.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tuma, Daniel  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William, I.  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; TITLE OF INVENTION: Acids Encoding the Same  
 ; FILE REFERENCE: 10466-14  
 ; CURRENT APPLICATION NUMBER: US/09/906,700  
 ; CURRENT FILING DATE: 2000-09-18  
 ; PRIOR APPLICATION NUMBER: PCT/US00/04414  
 ; PRIOR FILING DATE: 2000-02-22  
 ; PRIOR APPLICATION NUMBER: US 60/143,048  
 ; PRIOR FILING DATE: 1999-07-07  
 ; PRIOR APPLICATION NUMBER: US 60/145,698  
 ; PRIOR FILING DATE: 1999-07-26  
 ; PRIOR APPLICATION NUMBER: US 60/146,222  
 ; PRIOR FILING DATE: 1999-07-28  
 ; PRIOR APPLICATION NUMBER: PCT/US99/20594  
 ; PRIOR FILING DATE: 1999-09-08  
 ; PRIOR APPLICATION NUMBER: PCT/US99/20944  
 ; PRIOR FILING DATE: 1999-09-13  
 ; PRIOR APPLICATION NUMBER: PCT/US99/21090  
 ; PRIOR FILING DATE: 1999-09-15  
 ; PRIOR APPLICATION NUMBER: PCT/US99/21547  
 ; PRIOR FILING DATE: 1999-09-15  
 ; PRIOR APPLICATION NUMBER: PCT/US99/23089  
 ; PRIOR FILING DATE: 1999-10-05  
 ; PRIOR APPLICATION NUMBER: PCT/US99/28214  
 ; PRIOR FILING DATE: 1999-11-29  
 ; PRIOR APPLICATION NUMBER: PCT/US99/28313  
 ; PRIOR FILING DATE: 1999-11-30  
 ; PRIOR APPLICATION NUMBER: PCT/US99/28564  
 ; PRIOR FILING DATE: 1999-12-02  
 ; PRIOR APPLICATION NUMBER: PCT/US99/28565  
 ; PRIOR FILING DATE: 1999-12-02  
 ; PRIOR APPLICATION NUMBER: PCT/US99/30095  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: PCT/US99/30911  
 ; PRIOR FILING DATE: 1999-12-20  
 ; PRIOR APPLICATION NUMBER: PCT/US99/30999  
 ; PRIOR FILING DATE: 1999-12-20  
 ; PRIOR APPLICATION NUMBER: PCT/US00/00219  
 ; PRIOR FILING DATE: 2000-01-05  
 ; NUMBER OF SEQ ID NOS: 423  
 ; SEQ ID NO 245  
 ; LENGTH: 713  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapien  
 US-09-906-700-245

Query Match 12.8%; Score 425.5; DB 4; Length 713;  
 Best Local Similarity 25.2%; Pred. No. 1.4e-26;  
 Matches 182; Conservative 74; Mismatches 258; Indels 207; Gaps 26;

QY 1 MAPPLILLILASGAAA-----CPLPCVCO-----NLSESLTLCARGLLFV 42  
 Db 4 LVAPLLIAMIWAGATATVPVPMVHPVPCPCQACQIRPMYTPRSSYRBAITVDCNDLFLTAV 63  
 QY 43 PPNVDRRTVELRLADNFIQALGPPDFRNMGTGLVDLTLSRAITRIGARAGDLESLSLH 102  
 Db 64 PPALPACTQTLLLOSNSIVAVDQSELGIANLTLELDLSQNSFSFARDCCDFHALPQLLSLH 123  
 QY 103 LDGNRLVELGTSLRGVNLQHLILSGNQGRIAPGAFD----- 141

```

Db 124 LEENQTLRLBDHSPAGIASLOEVLNHNQYRIAPRAFSGLSNLRLHNSNLRRAIDSR 183
QY 142 -----DPLESLEDJD 151
Db 184 WFEMLPNLEILMIGKVKDAIDMNFRLPLANLSVLAGMNLREISDYALEGLQSLESLS 243
QY 152 LSYNNLRQVPMAGIAMPALHTLNDHNLIDALPPGAF-----190
Db 244 FYDNLARVPRALQVPGELKFDLKNPILQVPGDFPAMHLKELGLNMEELVSDK 303
QY 191 -----QLGQSRDLUTSN-RLATLADPLFS-----RGDAEASPAPLY 228
Db 304 FALVNPETLKLITNNPRLSFHPRAFHLPQWETLNNNLSALHQCVESLPLQOE 363
QY 229 LSFSGNPLHCNCELLMNR-----RLARPDDLETCASPPGLAGRYFMAVEGEFS--CEP 280
Db 364 VGHGNPFRCDCTIRMANNAITGRVRFLEPOS-TLCAEPDLORLPVREVPFRMTDCLP 422
QY 281 PLIAR-HTQRLWVLEGGRAITLRCRALGDPARTMHWVPPD-RLVGNSS-RADAFNGTL 336
Db 423 LISPSPFPPLQVASSGSWVLHCRALAEPEPEIYWTGALRLTPHAGRRVYVEGTL 482
QY 337 EIGVTAGDAGGTCTATNAGATARVELRYLALPHGNSSAEGGRPGPSDIPAASART 396
Db 483 ELRRVTAEEAGLYTCAQNLVGADITVSYV-----GRALDQPGD-----524
QY 397 AEGETLSESPAVQVETVTSGLVSW--GPRPADPVMWFQIQVNSSEDETLIYRIVPA 454
Db 525 -EGQ-LE-----LRVQETHYHLLISWYTPPNVSTNLTSSASSLSRGQATILARL-PR 577
QY 455 SSHEPLKLHVPADYDCLLALSPAGPSDLTATRLGCAHFTLPASPPLCHALQAHVL 514
Db 578 GTHSYNITRLQTEFWACIQV-----AFADAHTQLACVMARTEANS-CHR-----623
QY 515 GGLTIVAVG---GVVMAALVTFVALLVGRGA--GNGRLPLKLSHVQSGTNGGPPPTPK 569
Db 624 -----ALGDRPLI--AIIALAVLLIAGLAHLGTQ-PRK-----GVGGRPLDP 667
QY 570 A 570
Db 668 A 668

```

```

RESULT 5
US-09-903-603A-245
; Sequence 245, Application US/09903603A
; Patent No. 6767995
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gertschen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mathew, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey

```

```

; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: GNE.1618P2C12
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 245
; TYPE: PRT
; LENGTH: 713
; ORGANISM: Homo Sapien
; US-09-903-603A-245

```

```

Query Match 12.8%; Score 425.5; DB 4; Length 713;
Best Local Similarity 25.2%; Pred. No. 1,4e-26;
Matches 182; Conservative 74; Mismatches 258; Indels 207; Gaps 26;

```

```

QY 1 MAPLILLILLASGAA-----CPDPVCQ-----NLSSTLCAHRLTV 42
Db 4 LVAPLLAWVAGATATVPVPMHVPCPCQACQIRWYTPRSSYREATVDCNDLFTLV 63
QY 43 PENVDRRTVELELANFTIOALGPPDERMNTGLVDLTLSNATRLTGARFGDLESLSRH 102
Db 64 PPAIPAGTOTLLQNSIVRVQSELGTYLANTELDLSGNSPSDARDCCFHALLPOLLSLH 123
QY 103 LDGNRLVELGTSLRGPNVQLHLISGQGLRIAGAPD-----141
Db 124 LEENQTLRLBDHSPAGIASLOEVLNHNQYRIAPRAFSGLSNLRLHNSNLRRAIDSR 183
QY 142 -----DPLESLEDJD 151
Db 184 WFEMLPNLEILMIGKVKDAIDMNFRLPLANLSVLAGMNLREISDYALEGLQSLESLS 243
QY 152 LSYNNLRQVPMAGIAMPALHTLNDHNLIDALPPGAF-----190
Db 244 FYDNLARVPRALQVPGELKFDLKNPILQVPGDFPAMHLKELGLNMEELVSDK 303
QY 191 -----QLGQSRDLUTSN-RLATLADPLFS-----RGDAEASPAPLY 228

```

```

Db      304 FALVNLPELTKLIDITNNPRLSFHPRAFHLLPQMETIMLNNAALSAHQVTESLPNLOE 363
QY      229 LSSGNPLHNCNELLNR-----RLARPDLLETQSPGLAGRYFWAVPEGEFSC----- 278
Db      364 VGHGNPIRCDCVIRMANATGTRVRIEPOS-TLCAEPDLOQLPYREVPFRMTDCHLP 422
QY      281 PLIAR-HTORLWLEQORATLRCRALGDPAPYTMWVGPPDD-RLVGNSS--RARAFTNGTL 336
Db      423 LIPEPFPPELOVASGESWVLHCRALAEPELIYWTGPAQLRLTPHAGRRVYVEGTL 482
QY      337 EIGVTAGAGAGYTCIATNPAGEATYARVELVIALPHGNGSSAEGGRPEPSDIAASART 391
Db      483 ELRRVVAEEAGLYTCVQAQNLVGADITVSVV-----GRALLQPGRD----- 458
QY      397 ABEGTLESPPAQTVEVTITSGLVW--GPRPADPVMWFOIQYNSSEDETLIYRIPA 454
Db      525 -EGQG-LE-----LRVETHPYHLLSWTPPNTVSTNLTSASSLRGGQATLARI-PR 577
QY      455 SSHHFLKLHVPADYDLCIALSPAGPSDLTATRLGCAHFTLPASPLCHALQAVL 514
Db      578 GTHSYNTRLLQATEWACIQV-----AFADAHQOLACVMARTKENTS-CHR----- 623
QY      515 GGTITVAVG--GVLVAALLVFTVALIVRGGA--GNGRLPLKSHVQSOTNGPSETPK 569
Db      624 -----ALGDRPGLI--AIIALAVLLAAGLAHLGTGQ--PRK-----GVGGRRLPLP 667
QY      570 A 570
Db      668 A 668

```

## RESULT 6

```

US-09-063-950-2
; Sequence 2, Application US/09063950C
; Patent No. 6225085
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL LARG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OR INVENTION: THEREFOR
; FILE REFERENCE: MEI-019
; CURRENT APPLICATION NUMBER: US/09/063,950C
; CURRENT FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-063-950-2

```

```

Query Match      12.4%; Score 415; DB 3; Length 673;
Best Local Similarity 26.2%; Pred. No. 9.5e-26;
Matches 176; Conservative 69; Mismatches 280; Indels 148; Gaps 22;

QY      1 MAPPLLLILLASGAACPPLPCVQNTLSSTLCAHRLGLFVPPNVDRTVELERLADNFI 60
Db      8 LIRPLLLILLALGPVGQSPSGQC-----SQPQVFTAGTTPRDVDPDTGVLVFNNGI 64
QY      61 QALGPPDFRMTGLVDTLTSRNATIRIGARAFGLDESIRSLHLDGNRLVEIGTSLGPGV 120
Db      65 TMLDAGSFAGLPGLQQLDLDSQNTASLPSPGVQPLANISNLDLTANRLHEITNETFFGLR 124
QY      121 NLOHLILSGNQGRIFAGAPDFLESJEDLDLSYNNLRQVWAGIGAMPALHTNLNHL 180
Db      125 RLESLYIGKNIRIRITQGAFT-LDRLLLEKLQNELRALPPL--RLPRLILLDLSHNS 180
QY      181 IDALPP-----GAFAQLGQLSRLDITSN----- 203
Db      181 LIALEPGLDPTANVEALRLAGLGLQQLDEGLFSRLRNHLDDVDNQLERPVPIRGARG 240
QY      204 -----RLATLADPPLFS-----RGDAEASAPPL-----VLSFGSNPL 236
Db      241 LTRRLAGNTRIAQIRPEDIALGLAALQELDVSNLSIALPDLISGLPRLRLILAAARNPF 300

```

```

QY      237 HCNCELLMR-----RLARPDLLETQSPGLAGRYFWAVPEGEFSC----- 278
Db      301 NCVCPLSMFWGPMVRESHYTLASPEETR-CHFPKNAAGRLLELDYADFGCPATTTATAYP 359
QY      279 -----EPPLIARHTQRLWLEQORATLRCRALGDPAPYTMWVGPPDDRLVGNSSARAF 331
Db      360 TTRPVVEEPTLASSSIAPFTWLSPTPAT--EAPSPSTAPPTVGVPPQ-----PDQCP 410
QY      332 PNCITLIGVTGAGDAGYTCIATNPAGEATYARVELVIALPHGNGSSAEGGRPEPSDIA 391
Db      411 PSTCLNGGTCIHGTRHHLALCLC--PEFTGLVCE-----SQMGQTRPSPTPVIP 458
QY      392 SARTAEGEGLSEPAVQTEVTATSGLVSWGPRPADPVMWFOIQY--SSSEDETLIY 449
Db      459 RPRSL-----TLGIEP-----VSPSTSLAVGLQRYLQSSVQ-LRSLRLTYRLSGDPKRLVT 510
QY      450 RIVPASSHHFLKHLVPGADYDLCIALSPAGPSDLTATRLGCAHFTLPASPLCHAL 509
Db      511 LRLPASAETVYTLQLRPNATYSCVMPDLPGRVPEGEBA-----CGEANTPPVHSHNAP 565
QY      510 QAHVLGGLTV-----AVGVVAALLVFTVALIV-RGRG-----AGNGRLPLKLS 554
Db      566 VTQAREGNPLPLIAPLAAVLLAALAAVGAAYCYRRGRAMAAAADQXGVGPAGPLELE 625
QY      555 HVQSOTNGPSPPT 567
Db      626 GVKVPLEPGRKAT 638

```

## RESULT 7

```

US-09-907-794A-292
; Sequence 292, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Batton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvarolf, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gutney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28

```

Query Match	12.2%	Score 407.5;	DB 4;	Length 640;
Best Local Similarity	26.8%;	Pred. No. 3.7e-25;		
Matches 137; Conservative	66;	Mismatches 203;	Indels 105;	Gaps 16

QY	4	PLILLLIA-----SG---AAACPLPCYQCNLSLSLTLCAHRLGLFPFPPVNDRTVELR	54
Db	25	PLVLVLLATQLLVAGLVRAQTCBSVCSCN--QPSKVIYCRKNLEVPDGISTNTRLLN	82
QY	55	LADNFIALGDPDFRNMGLVDLTLSRNAITRIGARAFGDLESRLHDDGRVLVLTGT	114
Db	83	LHEHQIILIKYNSFEKLHLEILQLSRNRIETIEIGANNGLANNTLELPDRLLTTPNG	142
QY	115	SLRSPVNLQHLISGN-----QLGRIACGAPDFPLESED	149
Db	143	AFVYLSKLKELIMLNKNPPIESIPSAFNRIPELRRIDIGELKRLSYISGAFEG--LSNDRY	201
QY	150	LDLSYNNLRQYPMWAGIGAMPALHTLTNDHNLIDALPPGAFALQGLSL-----	198
Db	202	LNIAMCNLRIRP--NLNPLRLKLDDELDSGNLSAIRPSPFGIMHLQKTMIIQSIQYIE	259
QY	199	-----DLTSNRLATLAPDPLFSRGRDAEAPALVUSFSGNPLHCCNELIML	245
Db	260	RNAFDNLQSLVEINLAHNNT--TLIPHDLFTPLHLER-----IHLHNPNMNCNDIIML	312
QY	246	R--RLARPDLLEYCA---SPGGLAGRYFMVPEEGEFSCPEPLIARHTQRLMVLEGQAT	299
Db	313	SNWIKDWAPOSTACCACRNTPPNPKGRITGLDQNYFTCYAPVIVERPADLVNTBGMAAE	372
QY	300	LRGALGDPAPTMHWVGDDRLVGNSS---PARAPENGTEIEGVTGAGDAGGYTCIATNP	356
Db	373	LKCA--STSLTSVSWITPENGVTMTHGAYKRIALVSDLTINFTVNTVQDTIMYICMWSNS	431
QY	357	AGEATAVEELVIAL-----PHGQNSAEGRGPRPSDI-----AASART	395
Db	432	VGMTTASATINVTATTTPEFSYTVETVETMEPQDARKTDDNNVGPIPVVDWETTNVTT	491
QY	396	AABEGGLESEBPAAQVTEVTRATSGLVSWPG	426
Db	492	SLTPQSTRSTEKFTTIPVTDINSGI-----PG	518

```

1 RESULT 8
2 Sequence 292. Application US/09905125A
3 Patent No. 6664376
4
5 GENERAL INFORMATION:
6 APPLICANT: Genentech, Inc.
7 APPLICANT: Ashkenazi, Avi
8 APPLICANT: Botstein, David
9 APPLICANT: Desnovers, Luc
10 APPLICANT: Eaton, Dan L.
11 APPLICANT: Ferrara, Napoleone
12 APPLICANT: Filvaroff, Ellen
13 APPLICANT: Fong, Sherman
14 APPLICANT: Gao, Wei-Qiang
15 APPLICANT: Gerbers, Hanspeter
16 APPLICANT: Goddard, A.
17 APPLICANT: Godowski, Paul J.
18 APPLICANT: Grimaldi, Christopher J.
19 APPLICANT: Gurney, Austin L.
20 APPLICANT: Hillan, Kenneth, J.
21 APPLICANT: Kijavun, Ivar J.
22 APPLICANT: Mather, Jennie P.
23 APPLICANT: Pan, James
24 APPLICANT: Paoni, Nicholas F.
25 APPLICANT: Roy, Margaret Ann
26 APPLICANT: Stewart, Timothy A.
27 APPLICANT: Thomas, Daniel
28 APPLICANT: Williams, P. Mickey
29 APPLICANT: Wood, William, I.
30
31 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
32 FILE REFERENCE: 10466-14
33
34 CURRENT APPLICATION NUMBER: US/09/905,125A
35
36 CURRENT FILING DATE: 2001-07-12
37
38 PRIOR APPLICATION NUMBER: PCT/US00/04414
39
40 PRIOR FILING DATE: 2000-02-22
41
42 PRIOR APPLICATION NUMBER: US 60/143,048
43
44 PRIOR FILING DATE: 1999-07-07
45
46 PRIOR APPLICATION NUMBER: US 60/145,698
47
48 PRIOR FILING DATE: 1999-07-26
49
50 PRIOR APPLICATION NUMBER: US 60/146,222
51
52 PRIOR FILING DATE: 1999-07-28
53
54 PRIOR APPLICATION NUMBER: PCT/US99/20594
55
56 PRIOR FILING DATE: 1999-09-08
57
58 PRIOR APPLICATION NUMBER: PCT/US99/20944
59
60 PRIOR FILING DATE: 1999-09-13
61
62 PRIOR APPLICATION NUMBER: PCT/US99/21090
63
64 PRIOR FILING DATE: 1999-09-15
65
66 PRIOR APPLICATION NUMBER: PCT/US99/21547
67
68 PRIOR FILING DATE: 1999-09-15
69
70 PRIOR APPLICATION NUMBER: PCT/US99/23089
71
72 PRIOR FILING DATE: 1999-10-05
73
74 PRIOR APPLICATION NUMBER: PCT/US99/28214
75
76 PRIOR FILING DATE: 1999-11-29
77
78 PRIOR APPLICATION NUMBER: PCT/US99/28313
79
80 PRIOR FILING DATE: 1999-11-30
81
82 PRIOR APPLICATION NUMBER: PCT/US99/28564
83
84 PRIOR FILING DATE: 1999-12-02
85
86 PRIOR APPLICATION NUMBER: PCT/US99/28565
87
88 PRIOR FILING DATE: 1999-12-02
89
90 PRIOR APPLICATION NUMBER: PCT/US99/30095
91
92 PRIOR FILING DATE: 1999-12-16
93
94 PRIOR APPLICATION NUMBER: PCT/US99/30911
95
96 PRIOR FILING DATE: 1999-12-20
97
98 PRIOR APPLICATION NUMBER: PCT/US99/30999
99
100 PRIOR FILING DATE: 1999-12-20
101
102 PRIOR APPLICATION NUMBER: PCT/US00/00219
103
104 PRIOR FILING DATE: 2000-01-05
105
106 NUMBER OF SEQ ID NOS: 423
107
108 SEQ ID NO 292
109
110 LENGTH: 640

```

```
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-905-125A-292

Query Match
Best Local Similarity 26.8%; Score 407.5; DB 4; Length 640;
Matches 137; Conservative 66; Mismatches 203; Indels 105; Gaps 16;

QY 4 PLILLLIA-----SG---AAACPLPCVCONLSLSLSTLCARHGLFVPPNVDRIYELR 54
   |||:|||||
Db 25 PLIVVLLALQLLVAGVIGVRAQTCPVSCSN--QFSKVICVRKNLREVDPGISTNTRLIN 82
   |||:|||||
QY 55 LADNFIQALGPPPRNMTGVLDLTLSNATRIIGARAFGDLRSIHLDSNVLVEGTG 114
   |||:|||||
Db 83 LHENQIQTIVKNSKHLRHLFIQLDSRNHRTIETIGAFNGLANLNTLELFDNLTTPNG 142
   |||:|||||
QY 115 SLRGPNVLOHLISGN-----QLGRIAPGAFDDFLSLED 149
   |||:|||||
Db 143 AFVYLSKLKELWLRNNPTESI PSYAFNRIPSLRRLDGLKRLSYISEGAFEG-TSNLRY 201
   |||:|||||
QY 150 LDISYNNLRQVPMAGIGAMPALHTLNDHLIDALPGAFQQLGQLSRL----- 198
   |||:|||||
Db 202 LNLAMCNLRIRP--NLPLIKLDELDSGNHLSAIRGSPQGLMHLQKLMWISQIQVIE 259
   |||:|||||
QY 199 -----DLTSNRLATLAPDPLFSRGRDAEASPAVLVSFGNPLHNCCELLML 245
   |||:|||||
Db 260 RNAPFNLIQSIVEINLAHNNL-TLPHDLFTPLHLER-----IHLHNPWNCNDILML 312
   |||:|||||
QY 246 R---RLARPDDLETCA---SPGILAGRYFWAVBGEFSCEPPIIARHTORLWLEGORAT 299
   |||:|||||
Db 313 SWMIKQWAPENTACCAACNTPPNKKRYIGELDQNYTCAVPIVBPADLNTBGMAAE 372
   |||:|||||
QY 300 LRCRALGDPPTMHWGPDRLVNGNS--RARAFNGTIEIGVTGAGDAGTTCIATNP 356
   |||:|||||
Db 373 LKCRRA-STSLTYSWITPNGVTMTHGAYKRIAVLSGTLNFTNVTVDGMYTCWNSNS 431
   |||:|||||
QY 357 AGETARVELRYAL-----PHGNSSAEGGRPSDSI---AASART 395
   |||:|||||
Db 432 VGNTTATATNTVNTATTTPPSYSTVETVETMERPODARITDNNVGTTPVVDNETNVT 491
   |||:|||||
QY 396 AAEGETLSEBPAAVQVTEVATSGLVSWGPG 426
   |||:|||||
Db 492 SLTPQSTRSTEKFTTIVTDINSGL---PG 518
   |||:|||||

RESULT 9
US-09-902-775A-292
; Sequence 292, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Bolstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Nan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
```

```
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902, 775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 292
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-902-775A-292

Query Match
Best Local Similarity 12.2%; Score 407.5; DB 4; Length 640;
Matches 137; Conservative 66; Mismatches 203; Indels 105; Gaps 16;

QY 4 PLILLLIA-----SG---AAACPLPCVCONLSLSLSTLCARHGLFVPPNVDRIYELR 54
   |||:|||||
Db 25 PLIVVLLALQLLVAGVIGVRAQTCPVSCSN--QFSKVICVRKNLREVDPGISTNTRLIN 82
   |||:|||||
QY 55 LADNFIQALGPPPRNMTGVLDLTLSNATRIIGARAFGDLRSIHLDSNVLVEGTG 114
   |||:|||||
Db 83 LHENQIQTIVKNSKHLRHLFIQLDSRNHRTIETIGAFNGLANLNTLELFDNLTTPNG 142
   |||:|||||
QY 115 SLRGPNVLOHLISGN-----QLGRIAPGAFDDFLSLED 149
   |||:|||||
Db 143 AFVYLSKLKELWLRNNPTESI PSYAFNRIPSLRRLDGLKRLSYISEGAFEG-TSNLRY 201
   |||:|||||
QY 150 LDISYNNLRQVPMAGIGAMPALHTLNDHLIDALPGAFQQLGQLSRL----- 198
   |||:|||||
Db 202 LNLAMCNLRIRP--NLPLIKLDELDSGNHLSAIRGSPQGLMHLQKLMWISQIQVIE 259
   |||:|||||
QY 199 -----DLTSNRLATLAPDPLFSRGRDAEASPAVLVSFGNPLHNCCELLML 245
   |||:|||||
Db 260 RNAPFNLIQSIVEINLAHNNL-TLPHDLFTPLHLER-----IHLHNPWNCNDILML 312
   |||:|||||
QY 246 R---RLARPDDLETCA---SPGILAGRYFWAVBGEFSCEPPIIARHTORLWLEGORAT 299
   |||:|||||
```

Db 313 SWMTKMAPSNTACCAKRCNTPNVLKGRYIGELDONFTCYAPVIVPPADLNTBGMMAE 372  
Qy 300 LRCRALGDPAPTMHWGPDRLVGNSS---RARAFTNGTLEIGVTGAGAGGTYCTIATNP 356  
Db 373 LKCRRA-STSLTSVSMITPMTGVTMTHGAYKVRILAVLSDGTLNFTNVTVOOTGMVTCVWSNS 431  
Qy 357 AGEATARVELRYLAL-----PHGNSSAEGRPGPSDI---AASART 395  
Db 432 VGNTTASATLNTVTAATTPFSYSTVETVMEPSQDEARTDNNVGPFPVDMETNVT 491  
Qy 396 AAEGETLESEPAVOVTEVTATSGLVSWGPG 426  
Db 492 SLTPQSTSTKTEFTTIPVTDINSI-----PG 518

RESULT 10  
US-09-906-700-292  
Sequence 292, Application US/09906700  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Geritsen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kijavini, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/906,700  
CURRENT FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30

PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 292  
LENGTH: 640  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-09-906-700-292

Query Match 12.2%; Score 407.5; DB 4; Length 640;  
Best Local Similarity 26.8%; Pred. No. 3.7e-25;  
Matches 137; Conservative 66; Mismatches 203; Indels 105; Gaps 16;

Qy 4 PILLILLA-----SG---AAACPLPCVCONLSSTICARGLTFVPPNDRRTVEIR 54  
Db 25 PLIVVLLALQLLVAGLVRAQTCPSVCSN--QFSKVI CVRKNLEVPDGI STNRLLN 82  
Qy 55 LADNFIQALGPPDFRNMTGLVLTLSRNAITRIGARAFGDLSELRHLHDGRVLVLGIG 114  
Db 83 LHNQIQLIKVNSFHLRLHLEILQLSRNIRRTIEGAFGLNANLLELFDRLTTIPNG 142  
Qy 115 SLRGPNVLOHLISGN-----QGRIPGAFDFEESLD 149  
Db 143 ARVYLSKLELWLRNPNIESISYAFNRIPSRRLDGLKRLSTISGAFES-LSNLKX 201  
Qy 150 LDIYNNIRQVPMAGIGAMPALHTLNLDNLIDALPPGAFAQLGQLSRU----- 198  
Db 202 LNLAMCNIREIP--NLTPKIKDELDLSGNHLSAIRPGSFQGLMHLQKLMQISQIQTIE 259  
Qy 199 -----DTSNRLATLADPLFSRGRDAASAPVLVSSGNPLHNCCLMW 245  
Db 260 RNAFNLQSLVEINLAHNNL--TLPPDLFTPLHHER-----ILHNPNCNCDIWL 312  
Qy 246 R---RLARPDDLETC---SPPLAGRYPMAYBEFSGCEPLIARHTQRLWYLEGQRT 299  
Db 313 SWMTKMAPSNTACCAKRCNTPNVLKGRYIGELDONFTCYAPVIVPPADLNTBGMMAE 372  
Qy 300 LRCRALGDPAPTMHWGPDRLVGNSS---RARAFTNGTLEIGVTGAGAGGTYCTIATNP 356  
Db 373 LKCRRA-STSLTSVSMITPMTGVTMTHGAYKVRILAVLSDGTLNFTNVTVOOTGMVTCVWSNS 431  
Qy 357 AGEATARVELRYLAL-----PHGNSSAEGRPGPSDI---AASART 395  
Db 432 VGNTTASATLNTVTAATTPFSYSTVETVMEPSQDEARTDNNVGPFPVDMETNVT 491  
Qy 396 AAEGETLESEPAVOVTEVTATSGLVSWGPG 426  
Db 492 SLTPQSTSTKTEFTTIPVTDINSI-----PG 518

RESULT 11  
US-10-140-002-368  
Sequence 368, Application US/10140002  
Patent No. 6725730  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Geritsen, Mary E.  
APPLICANT: Goddard, Audrey

```

? APPLICANT: Godowski, Paul J.
? APPLICANT: Gurney,Austin L.
? APPLICANT: Sherwood,Steven
? APPLICANT: Smith,Victoria
? APPLICANT: Stewart,Timothy A.
? APPLICANT: Tunnas,Daniel
? APPLICANT: Watanabe,Colin K
? APPLICANT: Wood,William
? APPLICANT: Zhang,Zemin
? TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
? TITLE OF INVENTION: ACIDS ENCODING THE SAME
? FILE REFERENCE: P330R1C59
? CURRENT APPLICATION NUMBER: US/10/140,002
? CURRENT FILING DATE: 2002-05-06
? Prior Application removed - See Palm or File Wrapper
? NUMBER OF SEQ ID NOS: 550
? SEQ ID NO 368
? LENGTH: 640
? TYPE: PRT
? ORGANISM: Homo Sapien
?
US-10-140-002-368

```

	Query Match	12.2%	Score 407.5	DB 4	Length 640;
	Best Local Similarity	26.8%	Pred. No. 3.7e-25;		
	Matches 137;	Conservative	66;	Mismatches 20;	Indels 105; Gaps 16
QY	4	PLLLLLLA-----SG---AAACPLPCVCONLSBSLSTICAHKGLLFYPPVNDRTVELR	54		
Db	25	PLVLVLLALQLLVVAGLVRAQTCBVCSCSN--QBSKVIQVRKNLRREVBDGISTNTRLLN	82		
QY	55	LADNFIALGPDPDRNMGLVDLTISRALITRIGARAGDLESLSLHLDGRLVELGTG	114		
Db	83	LHENOQIQIKNSFKHLRHLLEILOSRNHRITRIBGAENGLANLNTLELFDRLTTIPNG	142		
QY	115	SLRGPVNLQHLILSGN-----QGRIRAGAPDFLESIED	149		
Db	143	AFVYLSLKKEIKWLKNRPISIESPAFNRIPISLRIDJLCEIKRLSTISEGAFEG-LSNLR	201		
QY	150	LIDLSYNNLRQVPMAGIGAMPALHTLNDLNDLIDALPGAFQALGQLSRU-----	198		
Db	202	LNLMCMNREIR- NLTPRLKLDLDELDSGNHLSAIRPSPGSLMHLQKLMIMIQSDIQVE	259		
QY	199	-----DITSRLATLADPLFSRCGRDAASPAVLVSFGSNPLHNCCELLWL	245		
Db	260	RNAFDNLQSLVEINLNAHNNL-TLPHDFTPLAHLEK-----IHLHNPMMNCDDIML	312		
QY	246	R---RLRPDDLEKTA---SPPGLAGRFMAVPEEESCEPPLIARHQRLVMLEGGQAT	299		
Db	313	SMWIKMAFSTACCARCNPENPKIGRATIGELDQVFTCYAAVIEPAPDLAVTGGMAE	372		
QY	300	LRCRALGDPAPFTMWCVPDDRLVGNSS---PARAFPNGLTIEGYAGAGAGYTCIATNP	356		
Db	373	LKGRF-SNLSLVSMSITPNGVIMHGAIKVRIAVLSDGLNTNTNTVYDQTGYTCMVNS	431		
QY	357	AGEATFARVELRVALL-----PHGNGSASBGRPGPSDI---AASAF	395		
Db	432	VGNITASATLNVTAATTPFSYFSTVIVETMEPSODEARTDDNNNGPTPVVDVMTETNVTT	491		
QY	396	AAEGBGLESEPAVQVTEVTASGLVSWGPG	426		
Db	492	SLNPOSTSTSEKFTTTPVDINSGL-----PG	518		

RESULT 12  
 US-09-903--6034-292  
 : Sequence 292, Application US/099036034  
 : Patent No. 6767995  
 : GENERAL INFORMATION:  
 : APPLICANT: Genentech, Inc.  
 : APPLICANT: Aghakhanzai, Avi  
 : APPLICANT: Botstein, David  
 : APPLICANT: Desnoyers, Luc  
 : APPLICANT: Eaton, Dan L.

APPLICANT: Periarra, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Geriltzen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OR INVENTION: Acids Encoding the Same  
FILE REFERENCE: GNE.1618P2C12  
CURRENT APPLICATION NUMBER: US/09/903,603A  
CURRENT FILING DATE: 2001-07-11  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 292  
LENGTH: 640  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-09-903-603A-292

```

Query Match      12.2%   Score 407.5; DB: 4; Length 640;
Beet Local Similarity 26.8%; Pred. No. 3,78-25;
Matches 137; Conservative 66; Mismatches 203; Indels 105; Gaps 16;

Qy      4  P L I L I I A ----- S G --- A A C P L P C V C N L S S I S T L C A H G L L V P P V D R Y E I R 54
      1  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      2  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      25  P L I V L I A L L I V A G L V R A O T C P S V C S C S N - O P S R Y I C R K N L R E P D G I S T N T R I L N 82

```





QY	66	PDERNMWGTGIVDITLTSNATTRGAAFCGLDLSLRSLHDNRILVELGSLRGVNLQHL	125
Db	50	P-----ANVTTLSTLSANKKITVLRKRAFADYQVYSLW-----	81
QY	126	ILSGNGLRIAPGAFDDPLESLEDIDLSYNNLRQVPMWAGICAMPALHTLIDNHLIDALP	185
Db	82	-LHNFNVRITVEPQAL-AVLSQIKNLDLSHNFISFSPWDLNMLALQILKXNHRIGSLP	139
QY	186	PGAFALQGLSLRDLTNSRLATLAPDPLFSPGRDAEAPFLVLSFGSNPLHCNCCELLMT	245
Db	140	RDLALGALPDRISIRINNRLFTLAQTF-----DALSLASHQILHNPHCCGGLVNL	192
QY	246	R-----RLARPDDLETCASPGGLAGRYFM-----AVEBEGFSCERPLARHTQRL	290
Db	193	QAWASTRVSLPEPDSI-ACASPPALQGVVYRPLPALPCADPSVHLSEAPLEAPGP--	249
QY	291	WVLEGGORALTRCALGDPAFTYHW-----YGPDRLVGN-----	324
Db	250	-LRAGLAFVLHCLADGHPTRRLQWOLQIPGTVLLEPVLVSGEDDVGAEGBEGSDGL	308
QY	325	-----SSRAVAFPNGLTEICVTGACDAGGYTCIATNPAGEATARY	364
Db	309	LTOQAQTPPAWPAWPAWPAWPAWPAWPAWPAWPAWPAWPAWPAWPAWPAWPAWPAW	368
QY	365	ELRVALLPHGQNSAAGCRPGPSDIAASKRTABEGCT-----	402
Db	369	RVAVVAATGPBKAFAPGAGGEDQAFTSEKSTAKRGNSVLSKPEGKIKGGLAKVSL	428
QY	403	--LESEP-----	407
Db	429	GEFETPEEDTSGGEBAEDQILADPAEBORCGNDPSRYVSNHAFNOSALKPHVFLGV	488
QY	408	-AVQVTE---VYATSGLVSMQBG-----RPADPYWM-----FOIQNSSED	444
Db	489	IADVVEREARVQUTPLAARMGPGGAGAPRPGRRPLRLYLCPAGGAAYQMSRYER	548
QY	445	ETLIVRIVPASHHFLKLHVPAVDVLCLLALSPAGPSDILTATR-----LLGCAH	496
Db	549	GVNMY-----MFRGIRPETNISTVCLALAGACHVQVYVFSIKKELPSLLIVAVSV	598
QY	497	F---STLP--ASPLCHALQAH	512
Db	599	FLVLVATVPILGAACCHLLAKH	620
RESULT 15			
US-09-907-794A-73			
Sequence 73, Application US/09907794A			
Patent No. 6635468			
GENERAL INFORMATION:			
APPLICANT: Genentech, Inc.			
APPLICANT: Aethkenazi, Avi			
APPLICANT: Botstein, David			
APPLICANT: Desnoyers, Luc			
APPLICANT: Eaton, Dan L.			
APPLICANT: Ferrara, Napoleone			
APPLICANT: Filvaroff, Ellen			
APPLICANT: Fong, Sherman			
APPLICANT: Gao, Wei-Qiang			
APPLICANT: Gerber, Hanspeter			
APPLICANT: Gerritsen, Mary E.			
APPLICANT: Goddard, A.			
APPLICANT: Godowski, Paul J.			
APPLICANT: Grimaldi, Christopher J.			
APPLICANT: Gurney, Austin L.			
APPLICANT: Hillan, Kenneth, J.			
APPLICANT: Kijavlin, Ivar J.			
APPLICANT: Mather, Jennie P.			
APPLICANT: Pan, James			
APPLICANT: Paoni, Nicholas F.			
APPLICANT: Roy, Margaret Ann			
APPLICANT: Stewart, Timothy A.			
APPLICANT: Tumas, Daniel			

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```
Db      323 LAVEPYAFRGLNYLRVLNVSQQLTTLEESVPHSVG-NLE-----TLIDSNPLACDC 375
Qy      241 ELIML-RRLARPD--DLFTCASPPGLAGRYFWAVPE---GEFSCEPLIA-RHTORLW 291
Db      376 RLIMVFRRRRLNPNRQOPTCATPEFVQKEFDPPDVLNPNYFTCRRARIRDKAQQVF 435
Qy      292 VLEGQRATLRCRALGDPAPTMHWVGPDRLVGNSSRAR--APPNGTLEIGVTGAGDAGY 349
Db      436 VDEGHTVQFVCRADGDPPEPALIWLSPRKHLVSAKSNGRLLVFPDGTLEVRVYQVQDNGTY 495
Qy      350 TCIATNPAGPATRKARVELRLVAL---PHGNSSAE--GGRPGPSDTAASARITAEGEGTL 403
Db      496 LCIAANAGGNDSPAHHLVRSYSPDWPHQPNKTFAFISNQPEGE-ANSTR-----ATV 548
Qy      404 ESEPAVOVTEVTATSGLVSW 423
Db      549 PPFEDIKTLLIATTWGFISF 568
```

Search completed: November 17, 2004, 22:41:41  
Job time : 43 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using SW model

Run on: November 17, 2004, 22:33:09 ; Search time 42 Seconds  
(without alignments)

1454.706 Million cell updates/sec

Title: US-10-071-879-10

Perfect score: 3336

Sequence: 1 MAPPLDLLILASGAACP...LGAGCRGVGSABRLSESV 635

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR 79;\*

1: pirl:\*

2: pirl:\*

3: pirl:\*

4: pirl:\*

## SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	372.5	11.2	421	2	hypothetical prote
2	343.5	10.3	707	2	neuronal leucine-r
3	308	9.2	1091	2	glial cell membran
4	292.5	8.8	789	2	hypothetical prote
5	292.5	8.8	1355	2	hypothetical prote
6	291.5	8.7	361	2	chondroaderin pre
7	287.5	8.6	1535	2	peroxidase - fru
8	279	8.4	1531	2	slit-1 protein hom
9	266.5	8.0	605	2	insulin-like growt
10	265.5	8.0	605	2	insulin-like growt
11	263.5	7.9	622	2	insulin-like growt
12	257.5	7.7	560	2	synleutin - human
13	253.5	7.6	1469	2	platelet membrane
14	253.5	7.6	1480	2	slit protein 1 pre
15	251.5	7.5	603	2	insulin-like growt
16	246	7.4	1523	2	MEG3 protein - ra
17	242	7.3	1025	2	insulin-like growt
18	239.5	7.2	626	1	secreted leucine-r
19	236.5	7.1	603	2	insulin-like growt
20	221	6.6	907	2	orphan G protein-c
21	213.5	6.4	382	2	proline- arginine
22	213	6.4	420	2	oncofetal tropob
23	213	6.4	662	2	garp precursor - h
24	212.5	6.4	575	2	hypothetical prote
25	211.5	6.3	1265	1	neural cell adhesi
26	210.5	6.3	680	2	hypothetical prote
27	207.5	6.2	312	1	leucine-rich alpha
28	207	6.2	375	2	fibromodulin precu
29	206.5	6.2	368	1	biglycan precursor

30	205.5	6.2	1328	2	T23007	hypothetical prote
31	204.5	6.1	369	2	S32793	biglycan precursor
32	203.5	6.1	369	2	S32559	biglycan precursor
33	203.5	6.1	369	2	S20811	proteoglycan I - m
34	203.5	6.1	458	2	T19941	hypothetical prote
35	202.5	6.1	907	2	JG0193	G protein-coupled
36	201	6.0	376	2	S55227	fibromodulin precu
37	198.5	6.0	536	2	A34901	lysine carboxypept
38	194.5	5.8	1097	2	A23943	Toll protein precu
39	192.5	5.8	1256	2	S60461	gene flightless-I
40	191.5	5.7	342	2	A46743	lumican precursor
41	191.5	5.7	738	2	T19938	hypothetical prote
42	191	5.7	1120	2	B86479	hypothetical prote
43	189.5	5.7	380	2	S71876	fibromodulin - chi
44	187	5.6	4302	2	A38971	polymeric kidney
45	186	5.6	338	2	S52284	lumicon, secretory

## ALIGNMENTS

### RESULT 1

T46266 Hypothetical protein DKFp761A179.1 - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004

C/Accession: T46266

R:Blum, H.; Baurerachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23034

A:Accession: T46266

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-421 <RA>

A:Cross-references: UNIPROT:Q9NT99; EMBL:AI137451

A:Experimental source: adult amygdala; clone DKFp761A179

C/Genetics:

A>Note: DKFp761A179.1

Query Match	11.2%	Score 372.5;	DB 2;	Length 421;
Best Local Similarity	29.0%;	Pred. No. 7.7e-17;		
Matches 115;	Conservative 41;	Mismatches 121;	Indels 119;	Gaps 14;
QY	67	DFRMTGLVDLTLSRNRATRIGARAFGDLESRLHDGNRLVELGTSLRGVNFQHLI	126	
DB	6	DIPNLTLV-----RLEELSGNRDLIRPSFGULSRKLW	44	
QY	127	LSGNQLGRIAPGAFDDLESLEDLISYNNLRQVPWAGIGAMPALHTLMDNLIDALPP	186	
DB	45	LMHAQVATERRAFDD-LKSLLELNLSHNNLSLP-----HDL-----	81	
QY	187	GMPAQQLGSLRLDLSNRRLATLAPDLPSRGDAEASPAIVLSFGNPLHNCCELLMR	246	
DB	82	--FTPLRLERVHL-----NHNPMHCNDVLMIS	108	
QY	247	---RLARPDLLETC---SPGLAGRYFAVEGEFCEPPLIARHTQLMLVLEGORATL	300	
DB	109	WMLKXVPSNTTCARCAHAPAGLKGKRYGELDQSHETCAPYIVERTPLDNTVEGMAEL	168	
QY	301	KCPALGDPAFTMHWGPPDRLVGNSS--RARAFTNGTLEIGVTAGAGAGYTCIATNPA	357	
DB	169	KCRT-GTSMTSYVWMLPNTGLMTHTGYSYRISVLHDGTLFTNVTVQDTGQYTCWVTNSA	227	
QY	358	GEATARVELRVLA---LPHGGSASAGCRPGPSDIAASARTAAEGBG-----TL	403	
DB	228	GNTTASATLNVSAVDPAAGTSGGGGGGGGGGGG-----GSGYTYFTTVTEVL	280	
QY	404	EEEPAVQVTEV-----TATSGLVSWGPRPAD	430	
DB	281	ETQPGREALQPRGTEKEPPPTTDGV--WGGRPGD	314	

RESULT 2

JC7763  
 C:Accession: A58532  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
 C/Accession: JC7763  
 R/Suzuki, Y.; Sato, N.; Tohyama, M.; Kanaka, A.; Takagi, T.  
 J. Biol. Chem. 271, 22522-22527, 1996  
 A/Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically in  
 R/Fukushima, K.; Matsuo, Y.; Kitahara, C.; Kuchino, Y.; Tsuda, H.  
 Biochem. Biophys. Res. Commun. 287, 257-263, 2001  
 A/Title: Rat neuronal leucine-rich repeat protein-3: Cloning and regulation of the gene  
 A/Reference number: JC7763; PMID:11549284  
 A/Contents: Fibrosarcoma cells  
 A/Accession: JC7763  
 A/Molecule type: mRNA  
 A/Residues: 1-707 <FUK>  
 A/Cross-references: UNIPROT:Q9RSY6; GB:AF291437  
 C/Comment: This protein, a new member of the neuronal leucine-rich repeat protein family  
 in protein-protein interaction and functions as a cell adhesion molecule or soluble ligand  
 C/Genetics:  
 A/Gene: nlr-3  
 C/Keywords: cell adhesion  
  
 Query Match 10.3%; Score 343.5; DB 2; Length 707;  
 Best Local Similarity 24.5%; Pred. No. 1.1e-14;  
 Matches 150; Conservative 77; Mismatches 183; Indels 203; Gaps 29;  
  
 QY 17 CPTPCVCO-----NLSSSLSTL-CAHKGILFVPPNVDRRTVELRLADNFIQAL-GPP 66  
 Db 29 CPQLCTCEIRPWPFRPSIYWEASTVDQNDGLNFPARLADYDILLQTNINARIHST 88  
 QY 67 DFR-NMTGLVDLTLSRNAITRIGARAFGDESLSLHLDGNRLVELGTSGLRGVNLQHL 125  
 Db 89 DFPVNLTG--LDLSQNNLSVTNINVOKMSQLSVYLENKLTLPKCKLYGSLNQEL 145  
 QY 126 ILISGNOGRAPGAF----- 140  
 Db 146 YVNNHLISLSPGAFVGLHNLRLHLNSNRLQMINSKFEPALPNLEILMGDPILRIKD 205  
 QY 141 -----DDFLESLEDLD-LSY--NNLRQVPMW----- 163  
 Db 206 MNQPPLKRLSVIAGINLTFEVPDDALVGLNLESISFYDNRLNKYQVALQKAVNLKFL 265  
 QY 164 -----GIGAMPALHTLINDHNLIDALP----- 185  
 Db 266 DLNKNPINRIRGDFSNMLHKLKELGINNMPEL--VSIDSLAVDNLPLRKITEATNNRSL 323  
 QY 186 ---PGAFQIGQSLRLDLSNRLATLAPDPLFSKGRAEASPAVLVSFGNPLHCCCL 242  
 Db 324 YIHENAFRLPKESLMLNSNALSYHGTI-----ESLPNIKEISHSNPIRCDCVI 376  
 QY 243 LMLR-----RLARPDLCTCASPGLAGR-----YFMAVPEGEFSCEPLIARHT--QR 289  
 Db 377 RWINMKNKTNRIFMPPDGL-FCVDPEPEFGQNVQVHRDMME--IC-LPLIAPESPSSI 431  
 QY 290 LMTLEGORATLRCALADPAFTMHWVGPD--DRIVGNSSRRAPF--PNGTLEI-GVTGAGD 345  
 Db 432 LDVEADSVSLHGCATAPQPEIYWIITPSSKRLPLNLRKIFYVHSGTLDIRGIT--PKE 490  
 QY 346 AGGTCTATPAGATARVELRLVALPHGSSASAGRPSPDSIAASARTAAEGEGLLES 405  
 Db 491 GGLYTCTAINLVG-----ADKSLIMYGVGFV-----PQD-----NNSGLN- 526  
 QY 406 EPAVQVTEVATSGLVSMGPRPADPYMWFQIQYNS--SEDETLIYRI-VPASSHHFL 461  
 Db 527 ---IKIRDIRANSVLVSM--KANSKILKSSVKWTATAVKTEDSQAAASARIPSDVKYNNL 580  
 QY 462 KHLVPGADYDLCL 474  
 Db 581 THLKPSLEYKICI 593  
  
 RESULT 3  
 A58532  
 glial cell membrane glycoprotein lig-1 precursor - mouse  
 C/Species: Mus musculus (house mouse)

C/Date: 11-Apr-1997 #sequence\_revision 11-Apr-1997 #text\_change 09-Jul-2004  
 C/Accession: A58532  
 C:Species: Rattus norvegicus (Norway rat)  
 C/Accession: A58532  
 R/Suzuki, Y.; Sato, N.; Tohyama, M.; Kanaka, A.; Takagi, T.  
 J. Biol. Chem. 271, 22522-22527, 1996  
 A/Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically in  
 R/Fukushima, K.; Matsuo, Y.; Kitahara, C.; Kuchino, Y.; Tsuda, H.  
 Biochem. Biophys. Res. Commun. 287, 257-263, 2001  
 A/Title: Rat neuronal leucine-rich repeat protein-3: Cloning and regulation of the gene  
 A/Reference number: A58532; PMID:11549284  
 A/Contents: Fibrosarcoma cells  
 A/Accession: A58532  
 A/Molecule type: mRNA  
 A/Residues: 1-1091 <SUZ>  
 A/Cross-references: UNIPROT:P70193; GB:D78572; NID:G1545806; PIDN:BA11416.1; PID:G15458  
 F/36-61/Domain: proteoglycan amino-terminal homology <PAH>  
 F/71-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>  
 F/95-117/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>  
 F/118-141/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>  
 F/142-165/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>  
 F/166-189/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>  
 F/191-213/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>  
 F/214-237/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>  
 F/238-261/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>  
 F/262-285/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR9>  
 F/286-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>  
 F/310-333/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>  
 F/334-357/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>  
 F/358-381/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>  
 F/385-408/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>  
 F/409-432/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>  
 F/440-485/Domain: proteoglycan carboxyl-terminal homology <PCH>  
  
 Query Match 9.2%; Score 308; DB 2; Length 1091;  
 Best Local Similarity 23.9%; Pred. No. 3.9e-12;  
 Matches 167; Conservative 63; Mismatches 282; Indels 188; Gaps 23;  
  
 QY 9 LLSGAACPLPCVCOQLSSSLSTLCAHKGIL-----FVPPNVDRRTVELRLADNFIQ 61  
 Db 179 ILSSGA-----FDGLSRSLTLRLSKNRITQLPYKAKLP---RLQLDLNRRIR 226  
 QY 62 ALGPPFRNMTGLVDLTLSRNAITRIGARAFGDESLSLHLDGNRLVELGTSGLRGVNL 121  
 Db 227 LIRGLTFQGLDSLEVLRLQNNISRLTDGAFWGLSKKHVHLHLYNSLVEYNSGSLYGLTA 266  
 QY 122 LQHLILSGNOGRAPGAFDPLESLEDLDLSYNNLRQVEMAGIGAMPALHTLINDHNL 181  
 Db 287 LHQLHLNNNSISRIQDQW-SFOCKLHELITSFNNILRLDEESLAEISLIRLSHNAI 345  
 QY 182 DALPPGAFQIGQSLRLDLSNRLA----- 206  
 Db 346 SHIAGAFKGLKSLRVLDLDHNEISGTIEDTSGAFTGLDNLKLTLPNGKIKSVAKRAFS 405  
 QY 207 -----TLAPPLFSGRDAEASPAFL-VLSFGNPLHCCCLWL-----RLARP 252  
 Db 406 GLESLHNLNGENALISVQDPAFAKMKLKEYLISSESFLCDQKLKPLPMWLMGRMLQAF 465  
 QY 253 DLCTCASPGLAGRYFMAVPEGEFSCE--PPLIARHTORLWLEGRATLRCALGDP 309  
 Db 466 VTATCAHPELSLKQSLFSVLPDSEFCDDPEKPOIITQPETMAVVGKIDIFTCSAASSS 525  
 QY 310 PTMHWGPDDRIV-----GNSSRAFAFPNGTLE-----IGVTGADAGYCTIATNPA 357  
 Db 526 SPMTFAKKDNEVLANNADMENFAHVRQDGEVNEYTTLHLRHVTFAHGRYQCLITNHF 585  
 QY 358 GEA---TARVELRVL---ALPHGSSASAGRPSPDSIAASARTAAEGEGLLESAPAVQ 410  
 Db 586 GSTYSHKARLTVAVLPSFTKIPH-----DIARIGTTARLCAATGHPNPQ 631  
 QY 411 VTEVATSGLVSMGPRPADPYMWFQIQYNSSEDETLIYRIVPASSHHFLKLVPGADY 470  
 Db 632 -----IAW-----QKDGTD-----PFAARERRM--HYMPDDDV 658  
 QY 471 DLCLIALSPAAGSDLTATRLGC---AHFTLPASPLCHALQAVNLGGTLTVAVGVL 526  
 Db 659 FTTTDVKIDMGVYSGTAQNSAGSVANSATLTVALETPLSLAVPLEDRV-----VTGEB-- 710

QY 527 VAALLVFTVALVGRGANGRL-----PLKLSHVQSGTNG----- 562  
DB 711 -----TVAFOCKATGSPRPRITWLGKGRPLSLTERHHPTPNQQLLVQNVWIDAGRY 763  
QY 563 -----GPSPTPKAHPPRPPRPPRQSGSLDLSGACGYA 557  
DB 764 TCMSNPLGTERAHSSQLSTLPTP--GCRKDGTTVGIFTTA 801

## RESULT 4

T28714  
hypothetical protein T21D12.9a - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C/Accession: T28714  
R/Moesener, J.  
submitted to the EMBL Data Library, August 1997  
A/Description: The sequence of C. elegans cosmid T21D12.  
A/Reference number: Z20514  
A/Accession: T28714  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-789 <WOB>  
A/Cross-references: EMBL:AF016687, PIDN:AA048096.1, GSPDB:GN00022, CESP:T21D12.9a  
A/Experimental source: strain Bristol N2; clone T21D12  
C/Genetics:  
A/Gene: CESP:T21D12.9a  
A/Map position: 4  
A/Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2

## Query Match

Best Local Similarity 8.8%; Score 292.5; DB 2; Length 789;  
Matches 136; Conservative 59; Mismatches 225; Indels 133; Gaps 18;

QY 51 VELRLADNFIQALGPDPFRMTGVLDTLSRNAITRIGARAGDLSLSLHLDGRLYE 110  
DB 106 VTILKLRNHITTLNOPSFSRLRKLKESLDLTRNMIRVRFLAFNQLPSLQNVSLARDVYR 165  
QY 111 LGTSLRGVNTLOHLILSGNQLGRIPAGFDDFLESLDLSYNNLRQVPMAGICAMPA 170  
DB 166 LDDGMFYACEGLKHLNLSNRQAVTEGMMFG-LTSLVLDLSYNOIQSFHSSWHTPK 224  
QY 171 LHTLNDHNLIDALPGAFPAQLGOLS-----RDLDTSNRLA 206  
DB 225 LKWLSTHSNRIOQLPSGSPFRVLKQLBELLSANSIDSLHKFALVGMSSLHKLDLSNTLA 284  
QY 207 TLAPDPL-----FSRG-----RDAASAPPLVLTSSGNP--LH----- 237  
DB 285 VCEVDGAVLYNTSMPLRLSLRFTNNQLRVLPKAFERFPALBELDLTDNPFIATHPEAFE 344  
QY 238 -----CNCELMLWR-----RLARPDDLTCASPPLAGRYFWAVEGEF 276  
DB 345 PLEIKRLVWNSSILDCQISWLASWIYRLKDKSIIAKCSYPPPLADLYVAITANL 404  
QY 277 SC--BPP--LIARHTQRLWVLEGQATLRCRALGDPAPTMW-----VGPDDRVLVG 323  
DB 405 TCHNDSPRAKIYRQPEVSTLIGEKARFTCNVYGASPLSIEWRVENGOPRVLVQDSATF 464  
QY 324 NSSRARAFNGTLEIGTVAG-----DAGGTCIATNPAG--EATARVRLALAP 372  
DB 465 LSTNRTAVVNGTFDERELAAABELLDNVAMTNSQCVARRNFSGDFSTHVLQYVQAP 524  
QY 373 HGNSSAEGRGPPSDIAASARTAAEGEGTLESEPAVQTEVATATGLVSWGGRPADVY 432  
DB 525 -----KTYTPEDMPLVLVGQTAKLCAATGTPRP--EIK 556  
QY 433 MNF-QIQNSSSEDETLIYRIVPASSHHFLKLHVPADYDLCIALSPAAGSDLTATRL 491  
DB 557 WAFEQIIPFAAEARRLY--VTENDHIIYMN--VTKEGQAVTCHATNVAAGTOASA-NL 611  
QY 492 LGCASFSTLPASP 504  
DB 612 IVFENFFHYPESP 624

## RESULT 5

T28715

hypothetical protein T21D12.9b - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C/Accession: T28715  
R/Moesener, J.  
submitted to the EMBL Data Library, August 1997  
A/Description: The sequence of C. elegans cosmid T21D12.  
A/Reference number: Z20514  
A/Accession: T28715  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1355 <WOB>  
A/Cross-references: EMBL:AF016687, PIDN:AA048095.1, GSPDB:GN00022, CESP:T21D12.9b  
A/Experimental source: strain Bristol N2; clone T21D12  
C/Genetics:  
A/Gene: CESP:T21D12.9b  
A/Map position: 4  
A/Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2; 786/3; 8

## Query Match

Best Local Similarity 8.8%; Score 292.5; DB 2; Length 1355;  
Matches 136; Conservative 59; Mismatches 225; Indels 133; Gaps 18;

QY 51 VELRLADNFIQALGPDPFRMTGVLDTLSRNAITRIGARAGDLSLSLHLDGRLYE 110  
DB 106 VTILKLRNHITTLNOPSFSRLRKLKESLDLTRNMIRVRFLAFNQLPSLQNVSLARDVYR 165  
QY 111 LGTSLRGVNTLOHLILSGNQLGRIPAGFDDFLESLDLSYNNLRQVPMAGICAMPA 170  
DB 166 LDDGMFYACEGLKHLNLSNRQAVTEGMMFG-LTSLVLDLSYNOIQSFHSSWHTPK 224  
QY 171 LHTLNDHNLIDALPGAFPAQLGOLS-----RDLDTSNRLA 206  
DB 225 LKWLSTHSNRIOQLPSGSPFRVLKQLBELLSANSIDSLHKFALVGMSSLHKLDLSNTLA 284  
QY 207 TLAPDPL-----FSRG-----RDAASAPPLVLTSSGNP--LH----- 237  
DB 285 VCEVDGAVLYNTSMPLRLSLRFTNNQLRVLPKAFERFPALBELDLTDNPFIATHPEAFE 344  
QY 238 -----CNCELMLWR-----RLARPDDLTCASPPLAGRYFWAVEGEF 276  
DB 345 PLEIKRLVWNSSILDCQISWLASWIYRLKDKSIIAKCSYPPPLADLYVAITANL 404  
QY 277 SC--BPP--LIARHTQRLWVLEGQATLRCRALGDPAPTMW-----VGPDDRVLVG 323  
DB 405 TCHNDSPRAKIYRQPEVSTLIGEKARFTCNVYGASPLSIEWRVENGOPRVLVQDSATF 464  
QY 324 NSSRARAFNGTLEIGTVAG-----DAGGTCIATNPAG--EATARVRLALAP 372  
DB 465 LSTNRTAVVNGTFDERELAAABELLDNVAMTNSQCVARRNFSGDFSTHVLQYVQAP 524  
QY 373 HGNSSAEGRGPPSDIAASARTAAEGEGTLESEPAVQTEVATATGLVSWGGRPADVY 432  
DB 525 -----KTYTPEDMPLVLVGQTAKLCAATGTPRP--EIK 556  
QY 433 MNF-QIQNSSSEDETLIYRIVPASSHHFLKLHVPADYDLCIALSPAAGSDLTATRL 491  
DB 557 WAFEQIIPFAAEARRLY--VTENDHIIYMN--VTKEGQAVTCHATNVAAGTOASA-NL 611  
QY 492 LGCASFSTLPASP 504  
DB 612 IVFENFFHYPESP 624

## RESULT 6

A53860

chondroaderin precursor - bovine  
N/Alternate names: 38k leucine-rich protein  
C/Species: Bos primigenius taurus (cattle)

C/Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 04-Apr-2004  
C/Accession: A53860  
R/Name: P.J.; Sommarin, Y.; Boynton, R.E.; Heinigard, D.  
J. Biol. Chem. 269, 21547-21554, 1994  
A/Title: The structure of a 38-kDa leucine-rich protein (chondroadherin) isolated from H  
A/Reference number: A53860; MUID:94342341; PMID:8063792  
A/Accession: A53860  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-361 <NEA>  
A/Cross-references: GB:U08018; NID:9470671; PIDN:AAA21330.1; PID:9470672  
C/Keywords: disulfide bond  
F;300-346/Domain: proteoglycan carboxyl-terminal homology <PCH>  
  
Query Match 8.7%; Score 291.5; DB 2; Length 361;  
Best Local Similarity 28.3%; Pred. No. 1.2e-11;  
Matches 100; Conservative 35; Mismatches 117; Indels 101; Gaps 10;  
  
QY 1 MAPPRLILLASG-----AAACPLPCVCQNLSESLTCAHRGLFVPPVDRRTVE 52  
D 1 MAPPRLILLASGSLGLASLALPALAACPOHCH--SDLQHYICDKVGLQK-PRVSEKTKL 57  
QY 53 LRLADNFIQALGPDPFRNMTGLVD-----LTLNRNATIRIG 88  
D 58 LNRNNFPLATNSFRAMNVLVSLHQHQIREVAAGAFRGKQLIYLYLSHNDIRVLR 117  
QY 89 ARAPGDLSELRSLHD-----GNRLVEIGTSLRGPVNLQH 124  
D 118 AGAFDDLTETLYLYLDHNKYTELRGLLSPLVLFILQNNNKIRELSGAFQACDLRW 177  
QY 125 LILSGNGLRIAPGAPDFLESLEDDLSTYNNLRQVPMAGIGAMPALHTNLNLDL 184  
D 178 LYSSENSLSQLGALDD-VENLAKFYLDNRQSLSYSAISKIRVVEELKSHNPLKSI 236  
QY 185 PGCAFAQLGQ-----LSRLDLSNRLATLAPDPLFSRGRD 219  
D 237 PDNAFGQFGRYLETMLDNTNLEKFSOGAFVGLVTLVHLENNRHLQLSPNPFDD- 292  
QY 220 AEASPAVLVSFGSNPLHNCCELLMLRR-----LARDDLETCSPPGLAGRY 267  
D 293 -----SLETLLTNPMWKTCTQLRMRDEAKTSRPD--ATCASPAPKFRGQH 338  
  
RESULT 7  
S46224  
peroxidasein - fruit fly (Drosophila sp.)  
C/Species: Drosophila sp.  
C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Oct-2000  
C/Accession: S46224  
R/Nelson, R.E.; Fessler, L.I.; Takagi, Y.; Blumberg, B.; Keene, D.R.; Olson, P.F.; Parke  
EMBO J. 13, 3438-3447, 1994  
A/Title: Peroxidasein, a novel enzyme-matrix protein of Drosophila development.  
A/Reference number: S46224; MUID:94341251; PMID:8062820  
A/Accession: S46224  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-1535 <NEU>  
A/Cross-references: GB:U11052; NID:9531384; PIDN:AAA61568.1; PID:9531385  
C/Superfamily: peroxidasein; myeloperoxidase homology; proteoglycan amino-terminal homolo  
F;19-44/Domain: proteoglycan amino-terminal homology <PAH4>  
F;661-1350/Domain: myeloperoxidase homology <MPX>  
  
Query Match 8.6%; Score 287.5; DB 2; Length 1535;  
Best Local Similarity 26.0%; Pred. No. 1.3e-10;  
Matches 108; Conservative 51; Mismatches 174; Indels 83; Gaps 12;  
  
QY 6 LILLASGAAA--CPLPCVCQNLSESLTCAHRGLFVPPNV--DRRTVELRADNFIQ 61  
D 13 LILLASGVSVCPCACTCLE-----RTVRCIRAKSAVPAKLPQDQTLDIR- 61  
QY 62 ALGPDPFRNMTGLVDLTLSRNAITRIGARAFGDLSELRSLHLDGNRLVELGTGSLRGPVN 121  
D 62 -----NHIEELPANAFSGIAQLTTLFLINDNELAVIQDGLNGLTA 101

C/Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 04-Apr-2004  
C/Accession: A53860  
R/Name: P.J.; Sommarin, Y.; Boynton, R.E.; Heinigard, D.  
J. Biol. Chem. 269, 21547-21554, 1994  
A/Title: The structure of a 38-kDa leucine-rich protein (chondroadherin) isolated from H  
A/Reference number: A53860; MUID:94342341; PMID:8063792  
A/Accession: A53860  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-361 <NEA>  
A/Cross-references: GB:U08018; NID:9470671; PIDN:AAA21330.1; PID:9470672  
C/Keywords: disulfide bond  
F;300-346/Domain: proteoglycan carboxyl-terminal homology <PCH>  
  
Query Match 8.7%; Score 291.5; DB 2; Length 361;  
Best Local Similarity 28.3%; Pred. No. 1.2e-11;  
Matches 100; Conservative 35; Mismatches 117; Indels 101; Gaps 10;  
  
QY 1 MAPPRLILLASG-----AAACPLPCVCQNLSESLTCAHRGLFVPPVDRRTVE 52  
D 1 MAPPRLILLASGSLGLASLALPALAACPOHCH--SDLQHYICDKVGLQK-PRVSEKTKL 57  
QY 53 LRLADNFIQALGPDPFRNMTGLVD-----LTLNRNATIRIG 88  
D 58 LNRNNFPLATNSFRAMNVLVSLHQHQIREVAAGAFRGKQLIYLYLSHNDIRVLR 117  
QY 89 ARAPGDLSELRSLHD-----GNRLVEIGTSLRGPVNLQH 124  
D 118 AGAFDDLTETLYLYLDHNKYTELRGLLSPLVLFILQNNNKIRELSGAFQACDLRW 177  
QY 125 LILSGNGLRIAPGAPDFLESLEDDLSTYNNLRQVPMAGIGAMPALHTNLNLDL 184  
D 178 LYSSENSLSQLGALDD-VENLAKFYLDNRQSLSYSAISKIRVVEELKSHNPLKSI 236  
QY 185 PGCAFAQLGQ-----LSRLDLSNRLATLAPDPLFSRGRD 219  
D 237 PDNAFGQFGRYLETMLDNTNLEKFSOGAFVGLVTLVHLENNRHLQLSPNPFDD- 292  
QY 220 AEASPAVLVSFGSNPLHNCCELLMLRR-----LARDDLETCSPPGLAGRY 267  
D 293 -----SLETLLTNPMWKTCTQLRMRDEAKTSRPD--ATCASPAPKFRGQH 338  
  
RESULT 8  
T42218  
slit-1 protein homolog - rat  
N/Alternate names: MEGR4 protein  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: T42218  
R/Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.  
Genomics 51, 27-34, 1998  
A/Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs  
A/Reference number: T42218; MUID:98360089; PMID:9633030  
A/Accession: T42218  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-1531 <NAK>  
A/Cross-references: UNIPROT:O88279; EMBL:AB011530; NID:93449289; PIDN:BAJ2460.1; PID:93  
A/Experimental source: strain Sprague-Dawley; Brain  
A/Genetics:  
A/Gene: MEGR4  
C/Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein r  
  
Query Match 8.4%; Score 279; DB 2; Length 1531;  
Best Local Similarity 22.0%; Pred. No. 4.5e-10;  
Matches 105; Conservative 43; Mismatches 123; Indels 206; Gaps 10;  
  
QY 3 PPLILLAS-----GAACPLPCVCQNLSESLTCAHRGLFVPPVDRRTVELRADN 58  
D 16 PELILLMAAAMRLGATACPALCTGTGYD-----CHGTGLQAIKNIIPNTERBELNGN 71  
QY 59 FIALGPDPFRNMTGLVDLTLSRNAITRIGARAFGDLSELRSLHLDGNRL----- 108  
D 72 NITRIHNDPAGIKOLRVLIQIMENOIGAVERGAFDMDKEERLRNLNQQLVPELFFON 131  
QY 109 -----VELGTGSL-----RGPVNLHLLISGNGLRIAPGAPDFLESLEDDLSTY 154  
D 132 NQALSRDLSENSLQAVPRARFGAIDDKQLDKNOISCBGAF-RALRGLEVLITLN 190  
QY 155 NNLRQVPMAGIGAMPALHT----- 173  
D 191 NNITIPVSEFNMPKLRTRLRSHNHLFCDCHLAWLSQWLRQRPFTGLFGCGSPASLRG 250  
QY 174 ----- 173  
D 251 LNVAYQSEFSGSGGEEAAQVPACTLSGSCPMSCSNGIYDCRGKGLTAIPANLPET 310  
QY 174 -----LNDHNLIDALPGAFQGLQSLRLDLSRLATLADPL- 213  
D 311 MTEIRLELNGIKSIIPGAFSPYRKLRIDLSNNQIAFIADACGLRSLSNLVLYGKIT 370  
QY 214 -FSRG-----RDAE-----ASPA 225

```

Db      371 DLPRGVFGGLVYTLQILLNANKINCIRPDAPFQDLNLSLSTLVNKGISIAKGTSTSLRA 430
Qy      226 PLVLSFSGNPIMHNCNCELLMLRLARPDDLET-----CASPGLAGRYFMVPEGEFSC 278
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      431 IQLTHLAQNPFICDCKMLKMLADFLKINPIETTTARCAAPRLANKIGQIKSKKFC 487

RESULT 9
insulin-like growth factor acid-labile chain - baboon
C/Species: Papio sp. (baboon)
C/Date: 17-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-May-1997
C/Accession: JCS5239
R:Delhanty, P.; Baxter, R.C.
Biochem. Biophys. Res. Commun. 227, 897-902, 1996
A/Title: The cloning and expression of the baboon acid-labile subunit of the insulin-like
A/Reference number: JCS5239; M0ID:97040714; PMID:8886027
A/Content: liver
A/Accession: JCS5239
A/Molecule type: mRNA
A/Residues: 1-605 <DEL>
C/Comment: This factor is structurally related to proinsulin and have insulin-like meta

Query Match      8.0%; Score 266.5; DB 2; Length 605;
Best Local Similarity 30.9%; Pred. No. 9.3e-10;
Matches 76; Conservative 26; Mismatches 93; Indels 51; Gaps 3;

Qy      16 ACPLPCVCQNLSE--SLSTLCARHGLLPVPNVDRRTVELRLADNFIOALGPDPFNNMG 73
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      40 ACATACACSYDDEVNELSVFCSSRNLTLPDGPGGTQALMLDNNLSVPPAFAFNLS 99
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      74 LVDLTISRNAITRIGARAFGDLSELSRLHLDGNRLVELGTG----- 114
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      100 LAFINLQGQGLSLEFQALGLENLCHLERNQKSLALGTFATTPALASGLSNRUS 159
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      115 -----SLRGPVNLQHILISGNQLGRIAPGAPDFLE 145
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      160 RLEDGFEBGLIMDLNMGNSLAVLPDAFRLGGLRELIVLAGNRLAVLQPALFSGLA 219
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      146 SLEDLDSYNNLRQVPMWAGAMPALHTLNDLNLIDALPRGAFQGLSRLDLSNRL 205
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      220 -LRELDLSRNALRAIKANVAVLPRLQKLYLDRNLIAAVALGAFGLKALRWLDLSHNRV 278
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      206 ATLAPD 211
          |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      279 AGLLD 284

RESULT 10
A41915
insulin-like growth factor-binding complex acid-labile chain precursor - human
N/Alternate names: Acid-labile subunit (ALS)
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C/Accession: A41915
R:Leong, S.R.; Baxter, R.C.; Camerato, T.; Dai, J.; Wood, W.I.
Mol. Endocrinol. 6, 870-876, 1992
A/Title: Structure and functional expression of the acid-labile subunit of the insulin-like
A/Reference number: A41915; M0ID:92357025; PMID:1379671
A/Accession: A41915
A/Status: preliminary
A/Molecule type: mRNA; protein
A/Residues: 1-605 <LEO>
A:Cross-references: UNIPROT:P35858; GB:M86826; NID:g184807; PIND:AAA36047.1; PID:g184808
A:Experimental source: liver
A/Note: sequence extracted from NCBI backbone (NCBI:P110171)
F/15-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
F/75-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
F/123-146/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
F/147-170/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
F/171-194/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
F/195-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>
F/219-242/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>

```

```

F/243-266/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>
F/267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR9>
F/291-314/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>
F/315-338/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
F/339-362/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
F/363-386/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
F/387-410/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
F/411-434/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
F/435-458/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
F/459-482/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>
F/483-506/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>
F/507-529/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR19>

Query Match      8.0%; Score 265.5; DB 2; Length 605;
Best Local Similarity 32.5%; Pred. No. 1.1e-09;
Matches 80; Conservative 26; Mismatches 89; Indels 51; Gaps 5;

Qy      16 ACPLPCVC--QNLSBSLSTLCARHGLLPVPNVDRRTVELRLADNFIOALGPDPFNNMT- 72
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      40 ACPAACVCSDDDADELSVFCSSRNLTLPDGPGGTQALMLDNNLSVPPAFAFNLS 99
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      73 -----GLVDLTISRNAITRIGARAFGDLSELSRLHLDGNRLV 109
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      100 LQFLINLQGQGLSLEFPQALGLENLCHLERNQKSLALGTFATTPALASGLSNRUS 159
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      110 EL-----GTGSL-----RGPVNLQHILISGNQLGRIAPGAPDFLE 145
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      160 RLEDGFEBGLSLMIDLNMGNSLAVLPDAFRLGGLSRLRELIVLAGNRLAVLQPALFSGLA 219
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      146 SLEDLDSYNNLRQVPMWAGAMPALHTLNDLNLIDALPRGAFQGLSRLDLSNRL 205
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      220 -LRELDLSRNALRAIKANVAVLPRLQKLYLDRNLIAAVALGAFGLKALRWLDLSHNRV 278
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      206 ATLAPD 211
          |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      279 AGLLD 284

RESULT 11
JC7973
synleutin - human
C/Species: Homo sapiens (man)
C/Date: 25-Aug-2003 #sequence_revision 25-Aug-2003 #text_change 15-Sep-2003
C/Accession: JC7973
R:Wang, W.; Yang, Y.; Li, L.; Shi, Y.
Biochem. Biophys. Res. Commun. 305, 981-988, 2003
A/Title: Synleutin, a novel leucine-rich repeat protein that increases the intensity o
A/Reference number: JC7973; PMID:12767927
A/Accession: JC7973
A/Molecule type: mRNA
A/Residues: 1-622 <WNA>
A:Cross-references: GB:AY280614
C/Comment: This protein that is a single span transmembrane leucine-rich repeat protei
e intensity of pleiotropic cytokine responses as an adhesion protein or a receptor.
C/Genetics:
A:Gene: slrn
A:Map position: 5q12.1
C/Keywords: cytokine; leucine-rich repeat; synleutin; transmembrane protein

Query Match      7.9%; Score 263.5; DB 2; Length 622;
Best Local Similarity 26.7%; Pred. No. 1.5e-09;
Matches 101; Conservative 35; Mismatches 121; Indels 121; Gaps 9;

Qy      6 LLLLLASGAACPPLPCVQNLSESLSTLCARHGLLPVPNVDRRT----- 50
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      21 LLLLLHKEKILGC--SSVCG-LCTGQINCRNLGLSTIPKPFESTVFLVLTGNNTIYNE 77
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      51 -----VELR-----LADNFIOALGPDPFNNMTGLVDL 77
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      78 SELTGHSVALVYLDNSNLTLYPKAFVQLRHLYFLFNNNFIKRLDPGIFKGLMLRLNL 137
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      78 TISRNAITRIGARAFGDLSELSRLHLDGNRLVELGTGSLRGPVNLQHILISGNQLGRAP 137
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```



Db 138 YLVQNVQSVFPRGVFNVLVSQVYLNLQRNLTVLGSQGFVGMALRIIDLSSNNNIIRISE 197  
 QY 138 GAPDFLESIEDLDLSTNNIRQVPWAGIGAMPALHTNLNDNLIDALPGAFQGLQLSR 197  
 Db 198 SGF-QHLENIACLTGLGNNILKVPNSNAFEVLKSLRRSLSHNPTEAIQPFKGLANLEY 256  
 QY 198 LDLTNNRLATLAPD-----PLFSRGRDAE-----221  
 Db 257 LLLKNSIRAVTRGCFSGINNLLKHLIISHNDLENLSDTSLLKNLYLKIDRRRIISID 316  
 QY 222 -----ASPAVLVSF-----SGNDLHONCELLMLRLARP 251  
 Db 317 NDFPENMGASLKIINLTFNNLTALHPVLKPLSLIHLQANSNMPECNCKLLGLRDWLAS 376  
 QY 252 DDLET---CASPEGLAGR 266  
 Db 377 SATLNTIYQNPSPMRGR 394

## RESULT 12

A60164  
 Platelet membrane glycoprotein V precursor - human  
 C/Species: Homo sapiens (man)  
 C/Date: 12-Jan-1993 #sequence, revision 24-Feb-1994 #text change 09-Jul-2004  
 C/Accession: A48030; A60164; A35483; C35483; A60432; A47507; S34329  
 R/Lanza, F.; Morales, M.; de la Salle, C.; Cazenave, J.P.; Clemetson, K.J.; Shimomura, T.  
 J. Biol. Chem. 268, 20801-20807, 1993  
 A/Title: Cloning and characterization of the gene encoding the human platelet glycoprotein V  
 A/Reference number: A48030; MUID:94012616; PMID:8407908  
 A/Accession: A48030  
 A/Molecule type: DNA  
 A/Residues: 1-560 <LA2>  
 A/Cross-references: UNIPROT:P40197; EMBL:Z23091; NID:G312501; PIDN:CAA0637.1; PID:G3125  
 R/Shimomura, T.; Fujimura, K.; Maehama, S.; Takemoto, M.; Oda, K.; Fujimoto, T.; Oyama, T.  
 Blood 75, 2349-2356, 1990  
 A/Title: Rapid purification and characterization of human platelet glycoprotein V: the a  
 A/Reference number: A60164; MUID:90275263; PMID:2350580  
 A/Accession: A60164  
 A/Molecule type: protein  
 A/Residues: 355-384, 'X', 386-390, 'X', 392-395, 'X', 397, 188-208, 'I', 210, 227-50, 'X', 52-53, 117-14  
 , 'X', 108, 'T', 161-72, 'TK', 75-77, 'V', 56-57, 'G', 479-487, 'X', 489-498, 'X', 500, 'X', 502-503, 'X'  
 R/Roch, G.J.; Church, T.A.; McWilliams, B.A.; Williams, S.A.  
 Biochem. Biophys. Res. Commun. 170, 153-161, 1990  
 A/Title: Human platelet glycoprotein V: a surface leucine-rich glycoprotein related to a  
 A/Reference number: A35483; MUID:90321220; PMID:2372284  
 A/Accession: A35483  
 A/Molecule type: protein  
 A/Residues: 145-166, 'I', 168-169, 'X', 171-172 <ROT>  
 A/Note: this proteolytic fragment was designated peptide M392  
 A/Accession: B35483  
 A/Molecule type: protein  
 A/Residues: 121-129, 'W', 131-135, 466-468, 'X', 470 <RO2>  
 A/Note: this material was designated peptide M393 but may contain two peptides  
 A/Accession: C35483  
 A/Molecule type: protein  
 A/Residues: 252-266, 'H', 268-272, 'X', 274-279, 'I', 281-284, 'I', 286 <RO3>  
 A/Note: this proteolytic fragment was designated peptide M401  
 R/Zafar, R.S.; Walz, D.A.  
 Thromb. Res. 53, 31-44, 1989  
 A/Title: Platelet membrane glycoprotein V: characterization of the thrombin-sensitive gl  
 A/Reference number: A60432; MUID:89162331; PMID:2922700  
 A/Accession: A60432  
 A/Molecule type: protein  
 A/Residues: 477-478, 'FX', 481-485, 'E', 487, 'V', 489-492, 'NQ', 495, 'E', 497-498 <ZAF>  
 R/Hickey, M.J.; Hagen, F.S.; Yagil, M.; Roch, G.J.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 8327-8331, 1993  
 A/Title: Human platelet glycoprotein V: characterization of the polypeptide and the rela  
 A/Reference number: A47507; MUID:93391348; PMID:7690959  
 A/Accession: A47507  
 A/Status: preliminary; translated from GB/EMBL/DBD  
 A/Molecule type: mRNA  
 A/Residues: 1-560 <RES>  
 A/Cross-references: GB:L11238; NID:g388759; PIDN:AAA03069.1; PID:g388760

C/Comment: This platelet membrane protein is a substrate for thrombin.  
 C/Comment: The amino end of the intact protein is blocked.  
 C/Comment: This protein is absent in Bernard-Soulier syndrome.  
 C/Genetics:  
 A/Gene: GDB:GPS  
 A/Cross-references: GDB:230236; OMIM:173511  
 A/Map position: 5pter-5qter  
 C/Keywords: blocked amino end; glycoprotein; platelet; tandem repeat; transmembrane prot

Query Match 7.7%; Score 257.5; DB 2; Length 560;  
 Best Local Similarity 25.7%; Pred. No. 3.3e-09;  
 Matches 144; Conservative 51; Mismatches 190; Indels 175; Gaps 22;

QY 6 LILLASGAACPLPCYC-QNISESISTCARGLLFPVPPVDR-----48  
 Db 10 VIGLRAQPPPCPCPKCFERDAACGCGVARISALGLPTNTLTHLLPQMGRCVLOSQS 69  
 QY 49 ---RTV--ELRLADNFIOALGPPDFRMTGLVDLTSRNAITRIGARAFGDLESRLHL 103  
 Db 70 FSGMTVLQRLMIDSHLSAVAPGTFSDILKTLRSLRNKITHLPGLLRKMLLEQFL 129  
 QY 104 DGNRLVELGTSLRGPNVLOHLISGNOLGRIAPGAFDPLESLEDLDSYNNIRQVPWA 163  
 Db 130 DHNALGIDQNMFOKLVNLQELALNONQDLFLPASLFTN-LENLKLLDLSGNMLTHLPKG 188  
 QY 164 GIGAMPALHTL-----NLDNLDAL-----PGARFQGLSRD 199  
 Db 189 LIGAQAKLERLLHSNRLVSLDSGLNSLGALETQFHRNHIRSIAPGARDRLPNLSLT 248  
 QY 200 LTSRLATLAPDPLFSRGRDAEASPAVLVSFGSNPLHNCNE-----LML- 245  
 Db 249 LSNNHIAFL-PSALFLHSNHL-----TLTLLENPL--AELPGVLFEGMGGLQELMIN 298  
 QY 246 -----RLARPDDLETCASPPGLAGRYFWAVPEGRS--CEPLIARHTQRLW 291  
 Db 299 RTQRLTLPAAFNNLSRLRYLGYTLSP-----RLSALPGARFQGLQELQVLAHSNGLT 352  
 QY 292 VLEGQATLRGRALGPAPFMHVGPPDRLVGSSSARAPNG-----354  
 Db 353 AL-----PDGLRLGILKRLQVSLRRNRRLRALPRLFNLSLESVQLDHNQ 398  
 QY 335 -TLEIGVTGA-----GDAGYTCIATNP--AG 358  
 Db 399 LETHPGVGFALRLTEVLLGHNSWCDCGLFGLMLRQHLGVG-----EEPPRCAG 453  
 QY 359 EATPARELVALLPHGNSABEGRPSPDIAASAPTAAREGGLT---ESEPAVQTEVT 415  
 Db 454 PG-AHAGLPLMALP-GGDACPCPPRGPFPAPADSSSEAPVHPALAPNSSEPPWMAQPV 511  
 QY 416 ATSGIVSWGPGRPADPVMP 435  
 Db 512 T-----GKGQDSHPWGF 524

## RESULT 13

B3665  
 slit protein 2 precursor - fruit fly (Drosophila melanogaster)  
 C/Species: Drosophila melanogaster  
 C/Date: 30-Apr-1991 #sequence, revision 30-Apr-1991 #text change 02-Aug-2002  
 R/Rotherberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.  
 Genes Dev. 4, 2169-2187, 1990  
 A/Title: slit: an extracellular protein necessary for development of midline glia and co  
 A/Reference number: A3665; MUID:91099665; PMID:2176656  
 A/Accession: B3665  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-1469 <ROT>  
 A/Cross-references: GB:X53959  
 C/Genetics:  
 A/Gene: FlyBase:sl  
 A/Cross-references: FlyBase:FBgn0003425  
 C/Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein r

F,66-91/Domain: proteoglycan amino-terminal homology <PAH1>  
 F,101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>  
 F,125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>  
 F,149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>  
 F,173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>  
 F,197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>  
 F,228-272/Domain: proteoglycan amino-terminal homology <PAH2>  
 F,288-313/Domain: proteoglycan amino-terminal homology <PAH3>  
 F,323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>  
 F,347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>  
 F,371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>  
 F,395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR9>  
 F,419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>  
 F,450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>  
 F,512-537/Domain: proteoglycan carboxyl-terminal homology <PCS1>  
 F,547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>  
 F,572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>  
 F,596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>  
 F,620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>  
 F,651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>  
 F,708-733/Domain: proteoglycan amino-terminal homology <PAH4>  
 F,743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>  
 F,767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>  
 F,846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>  
 F,1028-1061/Domain: EGF homology <EGF1>  
 F,1068-1099/Domain: EGF homology <EGF2>  
 F,1115-1148/Domain: EGF homology <EGF1>

Query Match 7.6%; Score 253.5; DB 2; Length 1469;  
 Best Local Similarity 20.8%; Pred. No. 1.9e-08;  
 Matches 104; Conservative 47; Mismatches 153; Indels 197; Gaps 11;

15 AACPLPCVCONLSSTLCAHGLLFVP---PNVDRRTVE----- 52  
 Db ARCPKVCSTGLNVD---CSHRGLTSVPRKISADVERLELQGNNTLVITYETDFQRLTKL 126  
 QY 53 --LRLADNFIQALGPPDFRNMGTGLVDLTLSRNAITRIGARAFGLDLSRLHLDGRNLYE 110  
 Db 127 RMLQLTNDQIHTIERNSPFDVLSLRLDLSNNVITTVGRRVFKAGQSLSLQDNNQITC 186  
 QY 111 LGTGSIRGPVNLQHLILSGNQ-----LGR----- 134  
 Db 187 LDEHAFKGLVELEILTINNNTNLSLPHNIFGSLGRRLARLSDNPFACDCHLSWLSRFLR 246  
 QY 135 ----- 134  
 Db 247 SATRLAPYRRCQSPQSLKQGNVADLHDFEKCSCGLTEHAPMEGGAENSCPHRCADGIY 306  
 QY 135 -----IAPGAFDDPLESIEDLISYNNLRQVPMAG 164  
 Db 307 DCKEKSILTSVPVTLPDQTTDVRLQGNFITELEPKSFSSP-RRLRRIDLNNNISRIADHA 365  
 QY 165 IGAMPALHTLNLIDHNLIDLPFGAFQOLQSLSDLTISNRATLADPLFSRGRDA----- 220  
 Db 366 LSGLKQLTTLVIVGKNIKQDLPSCGVFKGLSLRLLLNANEISCIKRDAR---RDLHSIS 421  
 QY 221 -----EASPAFLVLSFGNPLHNCCELTMLRRLARPDOLET---CAS 259  
 Db 422 LLSLYDNNTQSLANGFDLAKMSKVIHLAKNPICCNRLMLADYHKKPVIETSGARCS 481  
 QY 260 PEGLAGRYFWAVEPGEFSCPEPLIARHTQRLWLWLEGORATLRCRALGDPAFTMHWGPPD 319  
 Db 482 PKMHRRIEISREEFKCS-----WGELRMKLSGECRMDSDCPAMCHCGTIV 530  
 QY 320 RLVGNSR--ARAFTGTLLEI 338  
 Db 531 DCTGRRLKEIPRDIPLHTTEL 551

RESULT 14  
 A3665  
 silt protein 1 precursor - fruit fly (*Drosophila melanogaster*)  
 C/species: *Drosophila melanogaster*

C/Date: 30-Apr-1991 #sequence revision 30-Apr-1991 #text\_change 02-Aug-2002  
 C/Accession: A3665; A31640; S13523  
 R/Author: J.M.; Jacobs, J.R.; Goodman, C.S.; Aravanis-Tsakonas, S.  
 Genes Dev. 4, 2169-2187, 1990  
 A/Title: silt: an extracellular protein necessary for development of midline glia and C/Keywords: alternative splicing; growth factor  
 A/Reference number: A3665; PMID:9109665; PMID:2176636  
 A/Accession: A3665  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-1480 <R07>  
 A/Cross-references: GB:X53959; NID:98614; PIDN:CAA37910.1; PID:98615  
 R/Author: J.M.; Hartley, D.A.; Walthers, Z.; Aravanis-Tsakonas, S.  
 Cell 55, 1047-1059, 1988  
 A/Title: silt: An EGF-homologous locus of *D. melanogaster* involved in the development  
 A/Reference number: A31640; PMID:8907753; PMID:314436  
 A/Accession: A31640  
 A/Molecule type: DNA  
 A/Residues: 881-1182, 'G', 1185-1404, 'GT', 1463-1464, 'YHA', <R02>  
 A/Cross-references: GB:M23543; NID:9340939; PID:9514357  
 C/Genetics:  
 A/Genes: FlyBase:sil  
 A/Cross-references: FlyBase:FBgn0003425  
 A/Accession: 1351/3  
 C/Suprafamily: fruit fly silt protein; EGF homology; leucine-rich alpha-2-glycoprotein  
 C/Keywords: alternative splicing; growth factor  
 F,66-91/Domain: proteoglycan amino-terminal homology <PAH1>  
 F,101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>  
 F,125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>  
 F,149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>  
 F,173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>  
 F,197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>  
 F,228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>  
 F,288-313/Domain: proteoglycan amino-terminal homology <PAH2>  
 F,323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>  
 F,347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>  
 F,371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>  
 F,395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR9>  
 F,419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>  
 F,450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>  
 F,512-537/Domain: proteoglycan amino-terminal homology <PAH3>  
 F,547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>  
 F,572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>  
 F,596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>  
 F,620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>  
 F,651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>  
 F,708-733/Domain: proteoglycan amino-terminal homology <PAH4>  
 F,743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>  
 F,767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>  
 F,791-814/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>  
 F,815-838/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>  
 F,846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>  
 F,1028-1061/Domain: EGF homology <EGF1>  
 F,1068-1099/Domain: EGF homology <EGF2>  
 F,1115-1148/Domain: EGF homology <EGF1>

Query Match 7.6%; Score 253.5; DB 2; Length 1480;  
 Best Local Similarity 20.8%; Pred. No. 2e-08;  
 Matches 104; Conservative 47; Mismatches 153; Indels 197; Gaps 11;

15 AACPLPCVCONLSSTLCAHGLLFVP---PNVDRRTVE----- 52  
 Db ARCPKVCSTGLNVD---CSHRGLTSVPRKISADVERLELQGNNTLVITYETDFQRLTKL 126  
 QY 53 --LRLADNFIQALGPPDFRNMGTGLVDLTLSRNAITRIGARAFGLDLSRLHLDGRNLYE 110  
 Db 127 RMLQLTNDQIHTIERNSPFDVLSLRLDLSNNVITTVGRRVFKAGQSLSLQDNNQITC 186  
 QY 111 LGTGSIRGPVNLQHLILSGNQ-----LGR----- 134  
 Db 187 LDEHAFKGLVELEILTINNNTNLSLPHNIFGSLGRRLARLSDNPFACDCHLSWLSRFLR 246  
 QY 135 ----- 134

Db 247 SATRIAPYTRCQSPQLKGNVADLHDPEFKCSGLTEHAPMEGAENSCPHPCRCADGIV 306  
QY 135 -----IAPGAFDDLESDLDLSTNNLRQVPWAG 164  
Db 307 DCREKSLTSPVTLPPDITVRLQNPFITELPKPSFSF-RRLRRIDLNNNISRIADA 365  
QY 165 IGAMPALHTNLNDHLIDALPPGAFAQLGQLSRLDLSNRLATLAPDPLFSRGRDA----- 220  
Db 366 LSGIKQLTTLVLYGNKIKDIPSGVFKGLGSLRLLILNANFISCIKDAF----RDLHSL 421  
QY 221 -----EASPAPIVLSFGSNPLHNCCELLMRRLARPDDLET---CAS 259  
Db 422 LLSLYDNNISLANGTDAKMSKTVHLAKNPICDNLRLADYLHKNPLETSGARCES 481  
QY 260 PPLAGGYFNAVPGESCEPPLIARHTQRLWVLEGQATLRCALGDPAPTMHWGPDD 319  
Db 482 PKRMHRRIRISLREKFKCS-----WGELRMKLSGECRMDSDCPAMCHCEGTV 530  
QY 320 RLWGNSSR--ARAFPNGLTEI 338  
Db 531 DCTGRRLKEIIPDILPHTTEL 551

## RESULT 15

JC1282  
Insulin-like growth factor-binding protein acid labile chain precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C/Accession: JC1282  
R;Dal, J.; Baxter, R.C.  
Biochem. Biophys. Res. Commun. 188, 304-309, 1992  
A:Title: Molecular cloning of the acid-labile subunit of the rat insulin-like growth fac  
A:Reference number: JC1282; MUID:93038676; PMID:1384485  
A:Accession: JC1282  
A:Molecule type: mRNA  
A:Residues: 1-603 <DAI>  
A:Cross-references: UNIPROT:P35859; GB:S46785; NID:G258002; PIDN:AA823770.2; PID:G570593  
A:Experimental source: liver  
A:Note: the authors translated the codon AAG for residue 63 as Arg, AAA for residue 205  
F;1-27/Domain: signal sequence #status predicted <SIG>  
F;28-603/Product: insulin-like growth factor binding protein, acid labile chain #status  
F;267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 7.5%; Score 251.5; DB 2; Length 603;  
Best Local Similarity 28.4%; Pred. No. 8.8e-09;  
Matches 79; Conservative 38; Mismatches 90; Indels 71; Gaps 6;

QY 4 PLILLLLA-----SGAA-----CPLPCVC--QNLSESLTLCARGLLF 41  
Db 8 PALVVLAFWVALGPGCHLQGTDPGASADAEQPCPVACTSHDDYTDELGVFCSSKVLTH 67  
QY 42 VPPVNDRTVEIRLADNFICAGPDPFRNMTGLVDLTLSRNATIRIGARAFGDLESRL 101  
Db 68 LPDDIIVSTRALWDGNNLSISPSAAFQNLSSLDFTNLQGSWLRSLPQALGLQNLVYL 127  
QY 102 HLDGNRLVEIGTGLRGPVNLQHLILSGNQIRIAPGAFDDFLESLDLSYNNLRQVP 161  
Db 128 HLEKRNRLNLAVALGFTTTPSLASLSLSSNLGRLEBGLFOG-LSHWDNLNGWNSLVLP 186  
QY 162 -----MAGIGA-----MPALHT 173  
Db 187 DTWFQGLGNLHELVLAKNKLTYLQPALFCGIGELRELDLSRNALRSYKAVFVHLPRLOK 246  
QY 174 LNLIDHNLIDALPPGAFAQLGQLSRLDLSNRLATLAPD 211  
Db 247 LYLDRNLITAVAGAFILGMALRWLDLSHNRVAGLMED 284

Search completed: November 17, 2004, 22:40:56  
Job time : 46 secs